CHALMERS
UNIVERSITY OF TECHNOLOGY

Classification of epileptic seizures

Accelerometer based detection of hypermotor seizures

Master's thesis in Systems, Control and Mechatronics

Sofie Wållberg

Department of Electrical Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2017

Master's thesis 2017:EX049/2017

Classification of epileptic seizures

Accelerometer based detection of hypermotor seizures

SOFIE WÅLLBERG

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Electrical Engineering
Signal Processing and Biomedical Engineering
Chalmers University of Technology
Gothenburg, Sweden 2017

Classification of epileptic seizures
Accelerometer based detection of hypermotor seizures
SOFIE WÅLLBERG

Supervisor: Fredrik Ohlsson, RISE Acreo AB, Gothenburg
Examiner: Tomas McKelvey, Department of Electrical Engineering, Chalmers

Master's Thesis 2017:EX049/2017
Department of Electrical Engineering
Signal Processing and Biomedical Engineering
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Typeset in LaTeX
Gothenburg, Sweden 2017

Classification of epileptic seizures
Accelerometer based detection of hypermotor seizures
SOFIE WÅLLBERG
Department of Electrical Engineering
Chalmers University of Technology

Abstract

In order to provide objective data for physicians to evaluate, accelerometry can be used to monitor movements of patients suffering from epilepsy. The focus of this thesis is to evaluate data gathered at Sahlgrenska Academy with the purpose to build models which detects nocturnal hypermotor seizures (HMS). This thesis extends work from a previous study focused on detecting generalized tonic-clonic seizures (GTCS), and the ultimate goal is to combine the results to create multi-class classification algorithms to be implemented with wearable electronics. When compared to GTCS, HMS commonly have a larger variability with varying motoric responses and durations among the patients. To overcome the increased difficulty of HMS detection, two major approaches are considered in this thesis. The general approach where the classifiers are constructed using data from multiple patients and the patient specific approach where models are created for each patient using their individual data sets for both training and testing. A set of features from the previous work is considered and extended to incorporate common features used in literature regarding HMS detection. The feature space is used to train and evaluate the classification methods logistic regression, k-nearest neighbors (KNN), support vector machines (SVM), random forest and kernel density estimation (KDE). Additionally, pre-processed accelerometer data is used to train artificial neural networks (ANN) without first calculating features. It proved to be a difficult task to achieve perfect accuracy using the current data set, the listed methods and the implemented feature space. It was possible to achieve high sensitivity but at the cost of a large amount of false positives. By visual inspection of the accelerometer data and video recodings a set of subtle or atypical seizure could be identified. The results of the classification showed that these seizures were difficult to detect as expected from the inspectation, they were either too subtle or too atypical for accurate classification. Their similarity to normal non-seizure movements are most likely affecting the performance in a negative manner and several post-processing steps had to be considerd in order to reduce the amount to be within acceptable limits. Furthermore, the feature space was evaluated to find differences between GTCS and HMS, the results indicate that there are considerable differences and the feature space could serve as a basis for multi-class classification in future work.

Keywords: Machine learning, statistical analysis, epilepsy, hypermotor seizures, HMS, nocturnal, accelerometry, classification.

Acknowledgements

I would first like to thank my supervisor Fredrik Ohlsson, RISE Acreo AB, for his continuous support, advice and impeccable guidance. I would like to thank my examiner Tomas McKelvey, Chalmers University of Technology, for his time and advice during the project. I would also like to thank David Krýsl, Sahlgrenska Academy, for medical consultation and Dongni Johansson, Sahlgrenska Academy, for assistance in the literature study.

Sofie Wållberg, Gothenburg, June 2017

Contents

List of Figures **xiii**

List of Tables **xv**

1 Introduction **1**
1.1 Purpose . . . 2
1.2 Accelerometry . . . 2
1.2.1 Human motion . . . 2
1.2.2 Measurements . . . 2
1.3 Epilepsy and seizures . . . 3
1.3.1 GTCS . . . 3
1.3.1.1 GTCS patients . . . 3
1.3.2 HMS . . . 4
1.3.2.1 HMS patients . . . 6
1.4 Data limitations . . . 6
1.5 Aim . . . 7

2 Theory **9**
2.1 Data . . . 9
2.2 Features . . . 9
2.2.1 Vector Magnitude, Signal Magnitude Area, Accumulated Acceleration . . . 10
2.2.2 Root mean square, mean value . . . 10
2.2.3 Variance, Standard Deviation . . . 10
2.2.4 Correlation . . . 11
2.2.5 Entropy . . . 11
2.2.6 Jerk . . . 11
2.2.7 Spectral Edge Frequency . . . 12
2.2.8 Frequency bands, Frequency Peak . . . 12
2.3 Feature standardization . . . 12
2.4 Feature selection . . . 13
2.4.1 Forward selection . . . 13
2.4.2 Overlap integral . . . 13
2.5 Classification . . . 14
2.5.1 Kernel Density Estimation . . . 14
2.5.2 Logistic Regression . . . 15

2.5.3 K-Nearest Neighbors . . . 16
2.5.4 Random Forest . . . 17
2.5.5 Support Vector Machine . . . 18
2.5.6 Artificial Neural Network . . . 18
2.6 Evaluation . . . 20
2.6.1 Cross-validation . . . 20
2.6.2 Performance . . . 20
2.6.2.1 Binary classifier performance . . . 20
2.6.2.2 Classification performance . . . 21

3 Methods **23**
3.1 Algorithm structure . . . 23
3.2 Pre-processing . . . 25
3.2.1 Time segments . . . 25
3.3 Data reduction . . . 25
3.3.1 Activity based reduction . . . 26
3.3.1.1 Hour activity . . . 26
3.3.1.2 Quarter activity . . . 26
3.3.2 Kernel density estimation . . . 26
3.3.3 Subtle seizures . . . 27
3.3.4 Balancing dataset . . . 28
3.4 Features . . . 28
3.4.1 Feature selection . . . 28
3.4.1.1 Forward Selection . . . 28
3.4.1.2 Overlap Integral . . . 29
3.5 Classification . . . 29
3.5.1 Binary Classification . . . 30
3.5.1.1 Classification algorithms . . . 30
3.5.2 Post-processing . . . 31
3.5.2.1 Median filter . . . 31
3.5.2.2 Veto . . . 31
3.5.2.3 Decision epoch . . . 34
3.5.3 Performance . . . 34
3.6 Cross-validation . . . 34

4 Results **37**
4.1 Seizure statistics . . . 37
4.2 Low-activity build . . . 37
4.3 Features . . . 38
4.4 Subtle and atypical seizures . . . 41
4.5 Kernel Density Estimate . . . 43
4.5.1 General performance . . . 44
4.5.2 Patient specific performance . . . 46
4.6 Cross-validation . . . 50
4.6.1 Initial evaluation . . . 50
4.6.1.1 Data balancing method . . . 51
4.6.1.2 Classification parameters . . . 51

4.6.2 Post-processed example . . . 51
4.6.3 Binary performance . . . 53
4.6.4 Post-processed performance . . . 53
4.6.4.1 Patient 21 . . . 54
4.6.4.2 Patient 36 . . . 57
4.6.4.3 Patient 37 . . . 61
4.6.4.4 Patient 39 . . . 65
4.6.5 Summary . . . 69
4.7 Multi-class Features . . . 71

5 Discussion **73**
5.1 Data set . . . 73
5.2 Data reduction . . . 73
5.2.1 Low-activity build . . . 74
5.2.2 Balancing dataset . . . 74
5.3 Features . . . 75
5.4 Performance . . . 75
5.4.1 KDE . . . 75
5.4.2 Post-processed statistics . . . 75
5.5 General vs Patient specific approach . . . 76

6 Conclusion **77**

Bibliography **79**

A Appendix 1: HMS Seizure data **I**

B Appendix 2: HMS General result **III**
B.1 Randomly balanced training set . . . III
B.2 Cluster balanced training set . . . IX

C Appendix 3: HMS Patient 21 result **XVII**
C.1 Randomly balanced training set . . . XVII
C.2 Cluster balanced training set . . . XXIII

D Appendix 4: HMS Patient 36 result **XXXI**
D.1 Randomly balanced training set . . . XXXI
D.2 Cluster balanced training set . . . XXXVII

E Appendix 5: HMS Patient 37 result **XLV**
E.1 Randomly balanced training set . . . XLV
E.2 Cluster balanced training set . . . LI

F Appendix 6: HMS Patient 39 result **LIX**
F.1 Randomly balanced training set . . . LIX
F.2 Cluster balanced training set . . . LXV

G Appendix 7: GTCS Seizure data **LXXIII**

List of Figures

1.1 The motor behavior of a pre-processed GTC seizure. The response is measured using a sensor at each arm. 4
1.2 The motor behavior of a pre-processed hypermotor seizure. Sensor 2 have recorded a rythmic seizure manifestation. 5
1.3 The motor behavior of a pre-processed hypermotor seizure and a similar normal movement. The seizure is short and subtle. 5

2.1 Random forest with three trees 17
2.2 Schematics of a single hidden layer, feed-forward neural network. . . . 19

3.1 Flowchart describing the most significant parts of the classification process. 24
3.2 Flowchart describing the post-processing from the binary classification result to final classification performance. 32
3.3 The effect of the median filter and decision epoch of a binary classification result. Each second is classified, where a seizure classified second is represented by a peak and normal by a valley. 33

4.1 Left: Correlation of all HMS patients. Right: Entropy of all HMS patients. 39
4.2 Left: Jerk of all HMS patients. Right: Spectral edge frequency 95% of all HMS patients. 39
4.3 Left: Frequency peak of all HMS patients. Right: Vector magnitude of all HMS patients. 40
4.4 Left: Mean value of all HMS patients. Right: Standard deviation of all HMS patients. 40
4.5 Left: Entropy for all typical seizures. Right: Entropy for all subtle and atypical seizures. 41
4.6 The motor behavior of the fifth pre-processed hypermotor seizure of patient 37. Typical behavior. 42
4.7 The motor behavior of the sixteenth pre-processed hypermotor seizure of patient 37. The seizure is very subtle. 42
4.8 The motor behavior of the third pre-processed hypermotor seizure of patient 21. The seizure is similar to normal movements. 43
4.9 The motor behavior of the seventh pre-processed hypermotor seizure of patient 21. The signal is interpolated in some parts. 43

4.10 General KDE, estimate based on three patients and evaluated using the remaining patient. Left: Evaluated using patient 21. Right: Evaluated using patient 36. 45
4.11 General KDE, estimate based on three patients and evaluated using the remaining patient. Left: Evaluated using patient 37. Right: Evaluated using patient 39. 46
4.12 Patient specific KDE, estimate and evaluation using patient 21. . . . 48
4.13 Patient specific KDE, estimate and evaluation using patient 36. . . . 48
4.14 Patient specific KDE, estimate and evaluation using patient 37. . . . 49
4.15 Patient specific KDE, estimate and evaluation using patient 39. . . . 50
4.16 Left: Visualisation of the binary classification. Right: The corresponding post-processed classification. 52
4.17 The binary classification example with frequent false positives. 52
4.18 The post-processed example, signficicantly reduced amount of false positives. 53
4.19 Left: Entropy of patient 21. Right: RMS of patient 21. 72
4.20 Left: Spectral Edge Frequency 95 of patient 21. Right: Standard deviation of patient 21. 72

List of Tables

3.1 Combination table. 30

4.1 Low-activity build statistics, hour basis. 37
4.2 Low-activity build statistics, quarter basis. 38
4.3 General KDE performance 44
4.4 Patient specific KDE performance 47
4.5 False negatives patient 21. 54
4.6 Post-processed results for random forest, patient 21. 55
4.7 Post-processed results for KNN, patient 21. 55
4.8 Post-processed results for logistic regression, patient 21. 56
4.9 Post-processed results for SVM, patient 21. 56
4.10 Post-processed results for ANN, patient 21. 57
4.11 False negatives patient 36. 58
4.12 Post-processed results for random forest, patient 36. 58
4.13 Post-processed results for KNN, patient 36. 59
4.14 Post-processed results for logistic regression, patient 36. 59
4.15 Post-processed results for SVM, patient 36. 60
4.16 Post-processed results for ANN, patient 36. 61
4.17 False negatives patient 37. 62
4.18 Post-processed results for random forest, patient 37. 63
4.19 Post-processed results for KNN, patient 37. 63
4.20 Post-processed results for logistic regression, patient 37. 64
4.21 Post-processed results for SVM, patient 37. 64
4.22 Post-processed results for ANN, patient 37. 65
4.23 False negatives patient 39. 66
4.24 Post-processed results for random forest, patient 39. 66
4.25 Post-processed results for KNN, patient 39. 67
4.26 Post-processed results for logistic regression, patient 39. 67
4.27 Post-processed results for SVM, patient 39. 68
4.28 Post-processed results for ANN, patient 39. 69
4.29 Summary of post-processed results. 70
4.30 Rating of the summary of Table 4.29. 71

A.1 HMS statistics describing seizure onsets, duration, pre-processed duration and standard deviation. I

B.1 General cross-validation results for KNN. III

B.2 General cross-validation results for randf. IV
B.3 General cross-validation results for logistic regression. IV
B.4 General cross-validation results for SVM. V
B.5 General post-processed results for KNN. VI
B.6 General post-processed results for random forest. VII
B.7 General post-processed results for logistic regression. VII
B.8 General post-processed results SVM. VIII
B.9 General cross-validation results for KNN. IX
B.10 General cross-validation results for randf. X
B.11 General cross-validation results for logistic regression. X
B.12 General cross-validation results for SVM. XI
B.13 General post-processed results for KNN. XII
B.14 General post-processed results for random forest. XIII
B.15 General post-processed results for logistic regression. XIV
B.16 General post-processed results SVM. XIV

C.1 Cross-validation results for KNN, patient 21. XVII
C.2 Cross-validation results for randf, patient 21. XVIII
C.3 Cross-validation results for logistic regression, patient 21. XVIII
C.4 Cross-validation results for SVM, patient 21. XIX
C.5 Post-processed results for KNN, patient 21. XX
C.6 Post-processed results for random forest, patient 21. XXI
C.7 Post-processed results for logistic regression, patient 21. XXI
C.8 Post-processed results SVM, patient 21. XXII
C.9 Cross-validation results for KNN, patient 21. XXIII
C.10 Cross-validation results for randf, patient 21. XXIV
C.11 Cross-validation results for logistic regression, patient 21. XXV
C.12 Cross-validation results for SVM, patient 21. XXV
C.13 Post-processed results for KNN, patient 21. XXVI
C.14 Post-processed results for random forest, patient 21. XXVII
C.15 Post-processed results for logistic regression, patient 21. XXVIII
C.16 Post-processed results SVM, patient 21. XXVIII

D.1 Cross-validation results for KNN, patient 36. XXXI
D.2 Cross-validation results for randf, patient 36. XXXII
D.3 Cross-validation results for logistic regression, patient 36. XXXII
D.4 Cross-validation results for SVM, patient 36. XXXIII
D.5 Post-processed results for KNN, patient 36. XXXIV
D.6 Post-processed results for random forest, patient 36. XXXV
D.7 Post-processed results for logistic regression, patient 36. XXXV
D.8 Post-processed results SVM, patient 36. XXXVI
D.9 Cross-validation results for KNN, patient 36. XXXVII
D.10 Cross-validation results for randf, patient 36. XXXVIII
D.11 Cross-validation results for logistic regression, patient 36. XXXIX
D.12 Cross-validation results for SVM, patient 36. XXXIX
D.13 Post-processed results for KNN, patient 36. XL
D.14 Post-processed results for random forest, patient 36. XLI

D.15 Post-processed results for logistic regression, patient 36. XLII
D.16 Post-processed results SVM, patient 36. XLII

E.1 Cross-validation results for KNN, patient 37. XLV
E.2 Cross-validation results for randf, patient 37. XLVI
E.3 Cross-validation results for logistic regression, patient 37. XLVI
E.4 Cross-validation results for SVM, patient 37. XLVII
E.5 Post-processed results for KNN, patient 37. XLVIII
E.6 Post-processed results for random forest, patient 37. XLIX
E.7 Post-processed results for logistic regression, patient 37. XLIX
E.8 Post-processed results SVM, patient 37. L
E.9 Cross-validation results for KNN, patient 37. LI
E.10 Cross-validation results for randf, patient 37. LII
E.11 Cross-validation results for logistic regression, patient 37. LIII
E.12 Cross-validation results for SVM, patient 37. LIII
E.13 Post-processed results for KNN, patient 37. LIV
E.14 Post-processed results for random forest, patient 37. LV
E.15 Post-processed results for logistic regression, patient 37. LVI
E.16 Post-processed results SVM, patient 37. LVI

F.1 Cross-validation results for KNN, patient 39. LIX
F.2 Cross-validation results for randf, patient 39. LX
F.3 Cross-validation results for logistic regression, patient 39. LX
F.4 Cross-validation results for SVM, patient 39. LXI
F.5 Post-processed results for KNN, patient 39. LXII
F.6 Post-processed results for random forest, patient 39. LXIII
F.7 Post-processed results for logistic regression, patient 39. LXIII
F.8 Post-processed results SVM, patient 39. LXIV
F.9 Cross-validation results for KNN, patient 39. LXV
F.10 Cross-validation results for randf, patient 39. LXVI
F.11 Cross-validation results for logistic regression, patient 39. LXVII
F.12 Cross-validation results for SVM, patient 39. LXVII
F.13 Post-processed results for KNN, patient 39. LXVIII
F.14 Post-processed results for random forest, patient 39. LXIX
F.15 Post-processed results for logistic regression, patient 39. LXX
F.16 Post-processed results SVM, patient 39. LXX

G.1 GTCS statistics describing seizure onsets, duration, pre-processed duration and standard deviation. LXXIII

1

Introduction

Different types of epileptic seizures can be difficult to distinguish due to similar symptoms and behavior. Due to the similarities and the current necessity to rely on subjective information, misdiagnosis is common. In order to improve and simplify diagnosis machine learning algorithms could be used to process objective motion sensor data obtained from patients.

The work was performed in collaboration with RISE Acreo AB. Acreo is a partner in a project which aims to develop wearable electronics which can be used to simplify diagnosis and treatments of neurodegenerative diseases such as epilepsy. Besides Acreo, the partners in the project are Swedish Foundation for Strategic Research together with Sahlgrenska Academy at Gothenburg University, Swerea IVF, Swedish School of Textiles and MedTech West. The wearable electronics can be developed to measure motions and be used to collect data from patients outside clinical environments. The data collected can give an insight in how day-to-day activities are affected by epilepsy, how often the seizures occur and the character of the seizures. Continuously collecting data will however result in an enormous amount of data which is difficult for physicians to process. A successful implementation of machine learning algorithms can serve as a tool to physicians by extracting relevant data and help with reaching an accurate diagnosis. When treating the patients the wearable electronics can be used to evaluate which effect the treatment has by considering measured data rather than only relying on information given by the patients. Reports given by patients can be unreliable while data from implementing wearable electronics are expected to provide more objective data to be evaluated. Currently most clinical research is focused on GTCS detection while other types of epilepsy are less studied.

The Acreo project should however incorporate multi-class classification, ideally a model should be constructed to accurately identify and differentiate three classes of epileptic seizures; generalized tonic-clonic seizures (GTCS), psychogenic non-epileptic seizures (PNES) and hypermotor seizures (HMS). These seizure classes all have motor manifestations, enabling the possibility to evaluate accelerometer data. This thesis is focused on detection of HMS, machine learning models have been created to differentiate HMS from normal movements. The results and conclusions aquired will later be used in the Acreo project for multi-class classification and detection.

1.1 Purpose

The purpose of this thesis was to accurately identify and differentiate hypermotor seizures (HMS) from normal non-seizure movements using accelerometry while considering previous work done within the project. Ultimately the findings of this thesis will be used in the Acreo project for multi-class classification implemented to be used with wearable electronics.

1.2 Accelerometry

1.2.1 Human motion

Continuous evaluation of human motion is becoming more common, in most smartphones today there are health tracking possibilites such as pedometers. However, human movement can provide more information besides the number of steps taken each day. By evaluating movements it is possible to find indicators and evidence for disabilites, diseases and disorders. Accelerometer devices with their small size and low power consumption are ideal to record data continuously outside the clinical environment. Such devices can provide objective information for physicians to evaluate [1].

1.2.2 Measurements

The data is collected at a sample frequency of 50Hz using accelerometers, one placed on each arm and in some cases one additional sensor placed on the torso. Given that classification models are based on learning characteristics of data sets, the sensor placed on the torso is neglected as it is not available for all seizures.

The patients are restricted to the hospital environment during video-EEG monitoring and as a consequence the measurements are not collected during regular day-to-day behavior. If a motoric manifestation is considered harmful the physicians actively intervene and try to reduce the risk of physical harm. Clearly this may have a large impact on the motion sensor data gathered, when the physicians intervene to inhibit the motoric response some parts of the patients typical seizure behavior might not be accurately measured. This is however not a problem unique to the clinical setting, if the patient is monitored in the home environment the patient's next of kin might intervene in a similar manner.

Each set of measurements have been initially verified and analyzed by physicians and a log is provided which indicates seizure types and when the seizures occurs. The log and the different systems used for measuring are not perfectly synchronized and thus the motoric seizures must be identified in the accelerometer data. Since the number of patients is small this can be done manually by evaluating the raw data in proximity to the seizure time aquired from the log. In some cases when the seizure were not easily identified, physicians and EEG were consulted to determine the onset and duration.

1.3 Epilepsy and seizures

Epilepsy is a disease where recurring seizures are caused by abnormal neuronal activity in the brain. There are different types of epilepsy with varying symptoms which roughly depend on where in the brain the electrical event originates and how it spreads to other parts of the brain. There are a wide range of symptoms entirely depending on which areas of the brain are affected, such as speech impairment, muscle contraction and unprovoked feelings [2, 3].

In clinical research, seizures are defined in terms of EEG. However, since this thesis is performed using accelerometer data the focus is on motoric manifestations only. Seizure onset, duration and other variables used in the thesis are entirely referring to the motoric response and not the electrical activity in the brain. The motor onset is typically not coincident with the onset determined by EEG, there is a certain latency before the electric activity results in motor symptoms.

Using accelerometry to detect epilepsy have previously been used with promising results. Both GTCS and HMS have motor components which can be measured using accelerometers or gyroscopes. The use of accelerometers require the patient to be able to move their limbs freely, if the patient is restrained the measurements will not accurately describe the motoric behavior. The difficulties of the accelerometry approach is not to detect the seizures but to differentiate seizures from normal non-seizure movements. To reduce the amount of false alarms the patient could be equipped with a method to manually dismiss a seizure alarm if it is caused by normal movements [4].

1.3.1 GTCS

General tonic-clonic seizures is the most commonly known type of epilepsy. The seizure consists of two phases with different motoric symptoms. The tonic phase causes all muscles to contract at a high frequency with small magnitudes, the body stiffens and the person loses consciousness. When the clonic phase starts the contractions occurs with lower frequency but with higher magnitude, causing limbs to rythmically contract and relax. The two phases of GTCS have a very characteristic motoric response which does not vary much between individuals and usually lasts 1 to 3 minutes [5].

1.3.1.1 GTCS patients

The GTCS data used in this study is available from five patients, however only a fraction will be used for comparison with HMS. Figure 1.1 describes the behavior of a typical GTC seizure using acceleration with respect to time. The accelerometer data is pre-processed as described in Section 3.2. After the seizure some subtle movements are recorded which have significantly lower intensity and shorter duration.

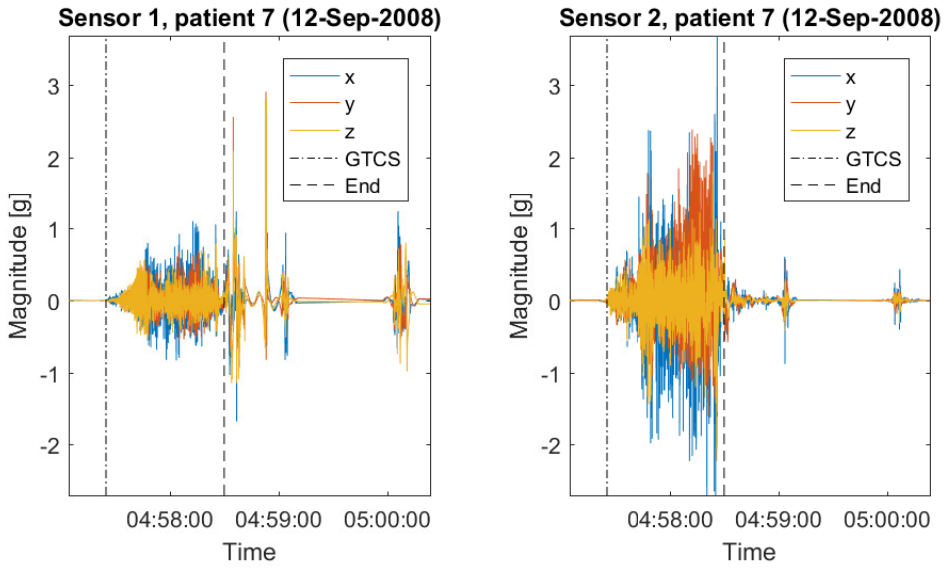

Figure 1.1: The motor behavior of a pre-processed GTC seizure. The response is measured using a sensor at each arm.

1.3.2 HMS

Hypermotor seizures consists of involuntary movements such as thrashing, kicking, boxing and rotation. In some cases there are severe cognitive distortions such as fear and panic. Unlike GTCS there is no typical pattern of movements, there is high variation in the motoric manifestations among patients. However, each patient often have a stereotypic seizure behavior. This implies that in order to perform a successful classification it might be beneficial to focus on individuals rather than using a general approach. The motoric manifestations usually last from a couple of seconds up to one minute [6].

In research related to HMS it is common to focus on nocturnal seizure monitoring since HMS typically occurs when the patient is asleep or at rest. Additonally, from a monitoring perspective, collecting data during the night is preferred. Epilepsy is determined by the use of video-EEG which might not be wireless, hence it is easier to gather measurements when the patient is in bed. Furthermore, nocturnal seizure monitoring is the setting from which the highest accuracy can be expected when using accelerometry. The short durations of the motoric manifestations can make it easy to mistake seizures for normal movements and vice versa. Naturally, monitoring performed when the patient is asleep will result in a data set consisting of movements which are much more subtle than if the monitoring had been performed during the day [7, 8].

A pre-processed HMS seizure is shown in Figure 1.2. If examined carefully, the x axis of sensor 2 includes a rythmic response, the patient appears to be hitting with one arm. This is verified by video monitoring. One additional example is shown in Figure 1.3, this seizure is very short and subtle. After the seizure there is a recorded

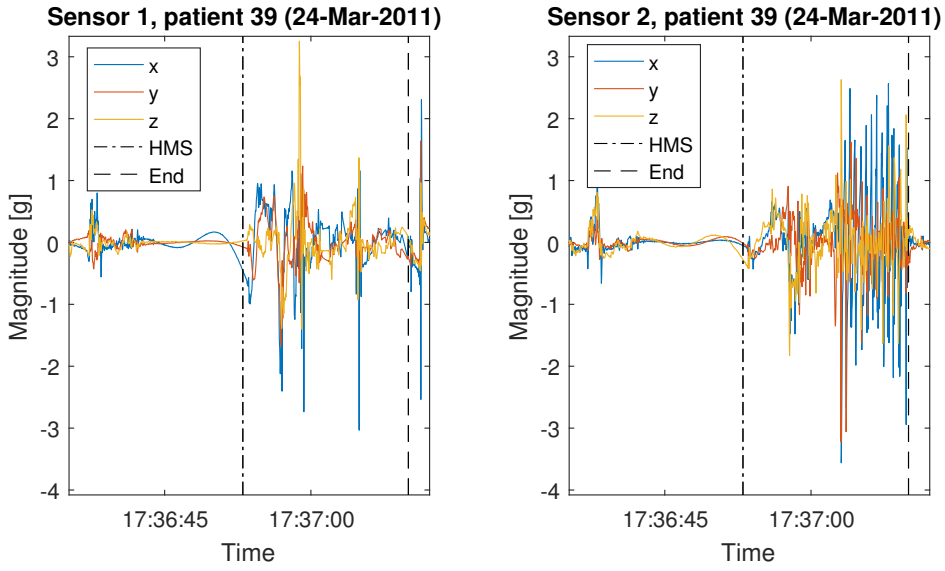

Figure 1.2: The motor behavior of a pre-processed hypermotor seizure. Sensor 2 have recorded a rythmic seizure manifestation.

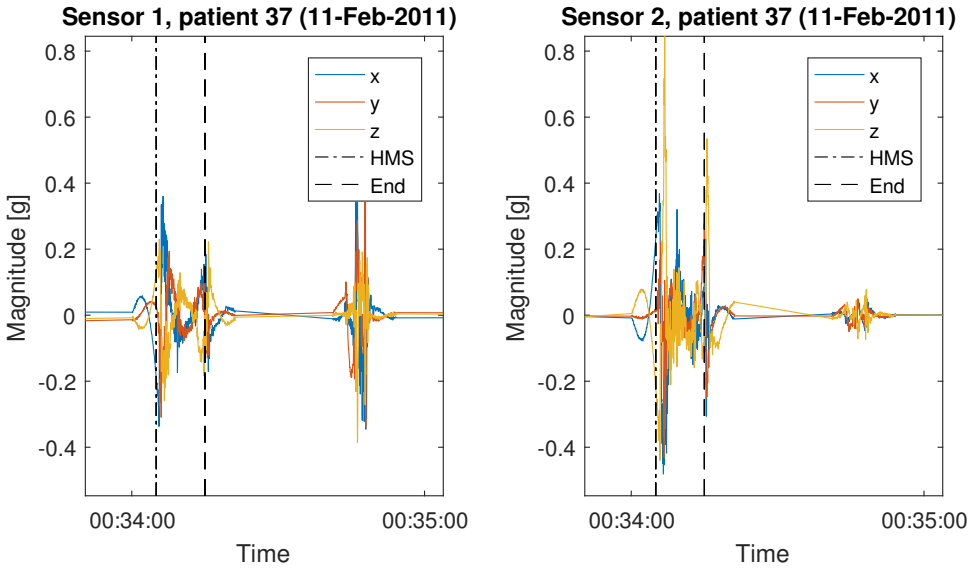

Figure 1.3: The motor behavior of a pre-processed hypermotor seizure and a similar normal movement. The seizure is short and subtle.

movement which is quite similar to the seizure. The most obvious difference is that the non-seizure movement seems to be focused on one arm while the seizure affects both. However, as in Figure 1.2, a seizure motoric manifestation is not neccessarily similar in both arms. By visual evaluation of graphs like these there is a wide range of motoric manifestations where most seizures seems to be quite similar to normal non-seizure movements, indicating that it might prove difficult to create an accurate classifier given the data set available.

1.3.2.1 HMS patients

There is HMS data available from four patients, patient 21, 36, 37 and 39. The patients have 8, 3, 24 and 2 seizures respectively. The onsets and durations of each seizure is listed in Table A.1 in the appendix.

Patient 21 have valid measurements from 62 hours, during which the patient had 8 HMS and 3 GTCS, where each GTCS occured a short time after a HMS. It is not uncommon for a seizure to evolve into a different type, depending on how the electricity spreads through the brain. When using the data from patient 21 for HMS classification the GTCS data is removed. The seizure durations are varying from 9 up to 32 seconds.

Patient 36 have valid measurements from 94 hours and a total of three seizures. The seizures have durations between 18 seconds and 31 seconds.

Patient 37 have valid measurements from 100 hours and 24 recorded seizures with motoric manifestations. The seizures have varying intensity, a few additional EEG recognized seizures had such a subtle motoric behavior that they are disregarded. The durations range from 6 to 22 seconds.

Patient 39 have valid measurements from 54 hours during which two HMS occured. Both seizures evolves into some other form of epileptic behavior which is disregarded. The seizures durations are 17 and 13 seconds respectively.

1.4 Data limitations

The data have been continuously gathered over several days which has resulted in a data set which mainly consist of regular movements and only a small fraction of seizures. Statistical analysis and machine learning algorithms are preferably performed using a large data set with many occurences of each class in order to estimate general distributions and similarities. The limited number of seizures introduces an unavoidable uncertainty in the results, there is no guarantee that the created models are optimal on a larger and more general scale.

1.5 Aim

The project is a continued research from a previous Master's thesis, "Machine learning for detection of epileptic seizures" [9]. Their focus was to identify GTCS and as a part of their project a set of features was evaluated and used with different types of machine learning algorithms. The aim of this work was to extend their research to identify nocturnal HMS instead. Initially, the previous set of features was evaluated on the new data to determine their relevance. A few additional features was implemented based on HMS related literature. The feature set was initially evaluated using probability based histograms. The features was used to create several types of classifiers with varying success, and additionally, the pre-processed data without feature extraction was evaluated using neural networks. The heavily imbalanced data set had to be considered and several methods was evaluated to reduce the impact it had on the classifiers. Furthermore, the classifiers proved to be prone to cause a large amount of false alarms which had to be dealt with. Lastly, the feature space was evaluated for multi-class classification purposes, to see if the features not only differentiated seizures from non-seizure but among different types of seizures as well.

The algorithms was trained and evaluated on the accelerometer data available. The data have been gathered from patients in hospital environment and does not fully represent everyday behavior. This is however the common practice in clinical research since it allows monitoring using video-EEG. Since epileptic seizures are defined by EEG, this type of monitoring is usually a requirement. The classification was performed offline however in a future stage classifiers can be trained offline and implemented to continuously evaluate data gathered from accelerometers.

The research questions considered in this work was

- How can machine learning algorithms be implemented to accurately detect and classify HMS?
- How can the specificity of the classifiers be increased?
- What are the differences between HMS and GTCS and which modifications to the previous work are required?
- How much data is needed to train the classifier?

2
Theory

2.1 Data

The data set consists of measurements from two accelerometer sensors, one located at each arm. Each sensor provides a set of three-dimensional data in terms of the acceleration in three orthogonal directions measured at a frequency of 50Hz, hence each measured second consists of 50 samples. The components of the sample i are denoted x_i, y_i and z_i.

2.2 Features

A set of features should be extracted to represent the characteristics of the accelerometer data, where each feature represents a specific aspect of the original data while omitting the rest. Representing the data as a set of features makes it possible to weigh the characteristics differently. For instance, if there is a specific component that is highly relevant to describe seizures it can be favored by the classifiers.

An evaluation of the feature space might reveal that there are some aspects which are non-essential for the classification process. Components which are basically equivalent in regular motion data and seizure data can be ignored as it does not improve a classifiers capabilitity to distinguish between said classes. Additionally, in order to perform accurately multi-class classification it is of importance to find characteristics which separates different types of seizures.

All features are calculated from the sampled data to result in one value for each measured second. Some features are calculated for each sensor while others are a combination of both sensors. Many features are calculated in combination with a Hamming window, w_i. The Hamming window ensures that each calculated second mostly depend on the centre of the considered samples, while the surroundings is given less importance. Some of the features are calculated with a window of two seconds with one second overlap to capture information over a longer period of time. In the previous thesis [9] some features for GTCS classification were calculated using a window of 10 seconds, however the duration of a HMS is usually much shorter than GTCS.

Most features considered in this work have previously been used to classify GTCS [9]. They are used for the HMS classification in this thesis since the ultimate goal

is to separate the GTCS from HMS. If these features can be used to differentiate HMS from normal non-seizure movements they should also be evaluated to see if it is possible to differentiate GTCS from HMS as well. Since many of these features are commonly used to evaluate motion data, they also occur in HMS research. Additonally, two new features Jerk and Spectral Edge Frequency are evaluated. Both features have been used in prior work focused on HMS classification.

2.2.1 Vector Magnitude, Signal Magnitude Area, Accumulated Acceleration

The magnitude of the signals can be evaluated in several ways, three different types are used to describe the general activity. Vector magnitude (2.1) is the mean of each sensor scaled using the window function. Signal magnitude area (2.2) is the mean of the absolute axis values. Accumulated acceleration (2.3) is the accumulated sum of the absolute values of each axis [9]. VM and SMA is calculated for each second in combination with a window function while AM is calculated over two seconds with one second overlap.

$$VM = \frac{\sum_i (\sqrt{x_i^2 + y_i^2 + z_i^2} \cdot w_i)}{\sum_i w_i} \tag{2.1}$$

$$SMA = \frac{\sum_i (\frac{|x_i| + |y_i| + |z_i|}{3} \cdot w_i)}{\sum_i w_i} \tag{2.2}$$

$$AM = \sum_{i=1}^{N} \sum_{j=1}^{i} (|x_i| + |y_i| + |z_i|) \tag{2.3}$$

2.2.2 Root mean square, mean value

Root mean square and the mean value is calculated for each sensor using the vector magnitude and a window of two seconds with one second overlap. These features have not only been used for GTCS detection, but for HMS as well [7, 8, 9].

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i^2 + y_i^2 + z_i^2)} \tag{2.4}$$

$$mean = \frac{1}{N} \sum_{i=1}^{N} (\sqrt{x_i^2 + y_i^2 + z_i^2}) \tag{2.5}$$

2.2.3 Variance, Standard Deviation

Variance and standard deviation are common statistical measurements used to describe variation within a dataset and have been used in prior epilepsy related research [7, 8, 10, 9]. The variance and standard deviation are calculated from the pre-processed data for each sensor using a window of two seconds with one second overlap.

2.2.4 Correlation

Correlation between sensors is calculated to evaluate similarities. A seizure with a rythmic motoric manifestation in both arms should result in significant correlation while normal movements are not neccessarily similar in both arms [8, 9]. The correlation is calculated by treating the sensor magnitude as random variables and calculating the mean for each second. Assuming N samples for each second, sensor magnitudes denoted X and Y and the sample mean μ, the correlation between the sensors is

$$CORR(X,Y) = \frac{1}{N}\sum_{i=1}^{N}(X-\mu_X)(Y-\mu_Y) \tag{2.6}$$

2.2.5 Entropy

Entropy is a measurement of information content and it is used to represent the general activity of measured movements. Given previous results, the feature is expected to provide a significant difference between seizure movements and non-seizure movements [9].

The entropy is calculated by evaluating the estimated probability density function of the acceleration. The probability density function $f(x)$ is estimated using 30 bins and a sliding window of two seconds with the MATLAB [11] function `histogram`. Given the probability density estimate $h_k(x)$ at bin k, the entropy is calculated

$$H(X) = -\sum_{k} f(x)\log f(x) \tag{2.7}$$

where

$$f(x) = \frac{h_k(x)}{\sum_k h_k(x)} \tag{2.8}$$

2.2.6 Jerk

The feature Jerk is introduced to represent how smooth and controlled movements are. Jerk denotes the rate of change of acceleration with respect to time and uncontrolled activity with short and fast movements results in higher jerk, hence it should be a suitable feature to describe uncontrolled seizure movements [8, 10].

Jerk is calculated for each measured second of each sensor. The first derivative is approximated using (2.9) for all directions and the absolute values are summarised in combination with a window function.

$$\dot{x}_i = \frac{x_{(i+1)} - x_i}{\Delta t} \tag{2.9}$$

$$Jerk = \frac{\sum_i (|\dot{x}_i| + |\dot{y}_i| + |\dot{z}_i|) \cdot w_i}{\sum_i w_i} \tag{2.10}$$

2.2.7 Spectral Edge Frequency

Spectral edge frequency is the frequency below which 80%, 90% or 95% of the total power of the signal is concentrated. It is used to capture and represent the general frequency content of the signal. It indicates whether the content consists of mainly low frequencies or contains frequencies from a wider spectrum [8]. A HMS seizure could for instance include the patient sitting up or rotating in the bed while at the same time rythmically thrashing with extremities. This type of seizure would result in frequencies from a wide spectrum and hence have a higher spectral edge frequency.

The spectral edge frequency is calculated using a sliding window of two seconds with one second overlap. Initially a periodogram is calculated using the MATLAB function `periodogram` [11], the periodogram is an estimate of the power spectral density of the signal. The function `bandpower` is used to estimate the integral of the power spectral density at the frequencies used by the periodogram function. The power in the frequency bands are evaluated to find the spectral edge frequency.

2.2.8 Frequency bands, Frequency Peak

The frequency content of the accelerometer signal can be evaluated by separating the content into frequency bands and estimating the signal power in the corresponding frequency ranges. A frequency band feature gives an estimate of the amount of energy in the signal and at which frequencies the activity occurs. Signal power in frequency bands and frequency peak have been used in prior research [7, 8, 9].

The feature is implemented using the MATLAB function `spectrogram` [11] which calculates the STFT of the signal at the frequencies of interest. To reduce the window effects and side lobes of the transform, the Hamming window w_i is used. The STFT is evaluated to find the energy content in each specified band, the energy of the signal is calculated as the sum of squared absolute values of the transform. The feature is calculated using a sliding window of two seconds.

Furthermore, the feature Frequency Peak is calculated by evaluating each second of the frequency bands to find at which frequency the energy content is at its maximum.

2.3 Feature standardization

In order to improve the ability to compare features, each feature is standardized. Standardization is important for most classification algorithms, by standardizing the features it ensures that the used classification algorithms focus on the feature distributions rather than the constant amplitudes. The feature space is standardized by assigning zero mean and unit variance for each feature. The mean μ_f and

standard deviation σ_f is calculated for each feature vector X_f in the training set. The mean and standard deviation of the training set will be used to standardize the test set as well to ensure similar scaling.

$$X_{norm,f} = \frac{X_f - \mu_f}{\sigma_f} \tag{2.11}$$

2.4 Feature selection

Several of the features describe similar properties which implies that all features might not be needed in order to capture the signal content. There might also be features which have similar non-seizure and seizure distributions, such features does not contribute to the classification process.

2.4.1 Forward selection

Forward selection is a simple greedy algorithm which can be used to select a subset of features by comparing the training error. The method can be used to reach a number of pre-defined features or by minimizing the total error. Forward selection is performed by iteratively adding the best choice of feature to the model. In the first iteration, the training error is calculated with only one feature, each available feature is evaluated and the best one is selected. In the next iteration the previously selected feature is iteratively combined with each of the remaining features and the training error is evaluated to choose the best combination. If the method is used to find a predefined number of features, the process is continued until the number of components have been chosen. If instead the goal is to minimize the total error irregardless of the number of features the algorithm should terminate when there is no feature left which reduces the training error.

Since forward selection does not consider all combinations the set of selected features is not neccessarily optimal, however it is still commonly used since testing each combination in the feature space is very time consuming unless the feature space is small [12].

2.4.2 Overlap integral

The overlap integral, or orbital overlap, is commonly used in quantum mechanics or chemistry to describe bonds between atoms. Simply put, a bond between atoms is formed when the orbit of its electrons overlap, if there is no overlap then no bond will be formed [13]. This concept can be applied to the feature space in this thesis as well. Classification algorithms highly depend on how the data is separated in the feature space. In the ideal case the data from different classes is separated entierly without overlap. This is usually not the case, but using the same idea it is logical to use the features which provide the most significant difference between seizure data

and regular motion data.

For each feature the probability density functions for the seizure data and normal data should be estimated. The overlap integral can then be calculated as a measurement of how well the seizure data is separated from the regular data. The overlap can be used to reduce the amount of features by selecting the features with smallest overlap. This method does not take into account the similarities between different features, for instance it is a possibility that the method selects both variance and standard deviation even though they are equivalent.

2.5 Classification

This section covers the classification algorithms used in the project. The goal of each method is to fit a model to describe how the set of classes y relates to the data set. The data set, denoted X, is a matrix of the feature space, where the columns corresponds to features and the rows correspond to time instances. The models are trained on a subset of X and tested on the remaining data, the models should make a prediction $\hat{y}$ for each time instance. Additionally, artificial neural network is used to directly evaluate the accelerometer data instead of using features.

2.5.1 Kernel Density Estimation

Kernel density estimation is a method to estimate probability density functions, f_X, based on a data set X. The method will be described in one dimension but it can be adapted to consider several dimensions. In the one dimensional case, the method iteratively places a kernel density function in each point of the training data set. After all kernel density functions have been placed, the method will iterate over all points once again to estimate the probability density function $\hat{f}_X(x)$ at each point x.

To achieve a smooth final estimate, the Parzen estimate (2.12) with weights is used. The equation describes the probability density estimate at point x, where each surrounding point is scaled with weigths determined by the kernel function K_σ. The weights decreases with distance, making the final estimation at each point the average contribution of the kernels in the neighborhood. Multiple kernels in the neighborhood of point x will make the point more probable, while few kernels in the neighborhood indicates a lower probability.

$$\hat{f}_X(x) = \frac{1}{N\sigma} \sum_{i=1}^{N} K_\sigma(x, x_i) \tag{2.12}$$

A common kernel choice is the Gaussian kernel. The Gaussian probability density function, $f_g(x|\mu, \sigma^2)$, with standard deviation σ and mean μ is

$$f_g(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{2.13}$$

Exchanging the kernel K_σ for the Gaussian kernel and considering the estimate at point x_i, it implies that the mean μ should equal x_i. In order to calculate the probability density estimate, each point x_i will be considered. The combination of (2.12) and (2.13) results in the estimate

$$\hat{f}_X(x) = \frac{1}{N\sqrt{2\pi\sigma^2}} \sum_{i=1}^{N} \exp\left(-\frac{(x-x_i)^2}{2\sigma^2}\right) \tag{2.14}$$

The kernel density estimate can be used for classification by either using outlier detection or by estimating a probability density function for each class. If each class is modeled by a probability density estimate, a test point can be classified by calculating the probability that it belongs to either class, choosing the most likely option. To make accurate estimations, multiple occurences of each class should be available for training [12].

In prior HMS classification research, multivariate kernel have been used for outlier detection of HMS [7, 8]. The probability density function is based on a feature space and estimated using only non-seizure movements. The estimate is evaluated to find a decision boundary which separates seizure movements from non-seizure movements. If a data point is outside of the decision boundary it is classified as a seizure. Ideally a decision boundary can be used to directly classify the data by searching for outliers, however if the features are not entirely separated there will be errors.

2.5.2 Logistic Regression

Logistic regression is as statistical method where the data is fitted to a linear model. In the binary case with only two classes, the logistic regression model will consist of only one linear function meant to separate the classes. Logistic regression differs from normal linear regression in one aspect, instead of being able to output a range of continuous values the logistic regression outputs the probability of belonging to a certain class. To describe a data set with K classes, the model consists of K-1 linear functions according to (2.15), where the syntax $Pr(G = k|X = x)$ denotes the probability that the data point x belongs to class k. The coefficients β are estimated during the training process.

$$log\frac{Pr(G=k|X=x)}{Pr(G=K|X=x)} = \beta_{k0} + \beta_k^T x, \quad k = 1, ..., K-1 \tag{2.15}$$

Since the goal is to assign the most probable class, the summarized probabilities should equal one:

$$Pr(G = 1|X = x) + Pr(G = 2|X = x) + ... + Pr(G = K|X = x) = 1 \tag{2.16}$$

Combining and rewriting (2.15) and (2.16) the probabilities for each class is as follows:

$$\begin{aligned} Pr(G = k|X = x) &= \frac{\exp(\beta_{k0} + \beta_k^T x)}{1 + \sum_{l=1}^{K-1} exp(\beta_{l0} + \beta_l^T x)}, \quad k = 1, ..., K-1 \\ Pr(G = K|X = x) &= \frac{1}{1 + \sum_{l=1}^{K-1} \exp(\beta_{l0} + \beta_l^T x)} \end{aligned} \tag{2.17}$$

The model is trained by maximizing the log-likelihood which depends on the coefficient vector β. In the binary case with only two classes, each predicted output $\hat{y}$ should be either 0 or 1. A training set of size N, known outputs y and the probabilites above results in the log-likelihood

$$l(\beta) = \sum_{i=1}^{N} (y_i \beta^T x_i - log(1 + \exp(\beta^T x_i)) \tag{2.18}$$

which is maximized to estimate β. For it to be maximized the derivatives of l must equal zero. To solve the system of equations and find the optimal values of β, the Newton-Raphson algorithm is used which iteratively searches for the optimal coefficients by evaluating the second derivatives of the system [12].

2.5.3 K-Nearest Neighbors

K-Nearest Neighbors is a classification algorithm where each point is classified by using the majority vote of the neighborhood. When performing classification of a data point x, the algorithm evaluates the euclidian distances in the feature space to find K data points which are closest in distance to the original point. The point is then assigned the most common class in the neighborhood. The method is defined as (2.19), where the predicted class $\hat{y}$ is based on the known classes of the training set, y_i, in the neighborhood N_K of x.

$$\hat{y}(x) = \frac{1}{K} \sum_{x_i \in N_K(x)} y_i \tag{2.19}$$

The only parameter in the algorithm is how many neighbors should be considered. Few neighbors usually leads to a noisy decision boundary where outliers have a large impact on the classification while a large number of neighbors leads to a smoother decision boundary but less precise where outliers have less impact [12].

2.5.4 Random Forest

Random forest is most easily explained by a figure. An example with three trees is shown in Figure 2.1, where the point x is classified as $\hat{y}$. During the training process, a random bootstrap sample is drawn from the training feature space for each tree to be trained. From the sample, a random subset of features is chosen and evaluated to find the best split-point among the features. The node is divided into two daughter nodes according to the split-point. The concept of split-points can be interpreted as questions, in the figure one split-point is referred to as q_1, and corresponds to a question which can be answered with yes or no. For instance, a question could be "is the vector magnitude larger than 2?" and depending on the answer different subtrees are entered. Each tree is trained iteratively by evaluating a random subset of features, finding the best split-points and split each node accordlingly. This is repeated until the stopping critera have been reached. There are several stopping criterias which can be considered. The most basic approach is to stop splitting nodes when the tree have reached a predefined size. One common criteria considers the response of each node, if all samples in the node have the same response it makes no sense to split the node further as it will not change the result, hence the splitting of that particular node will be stopped.

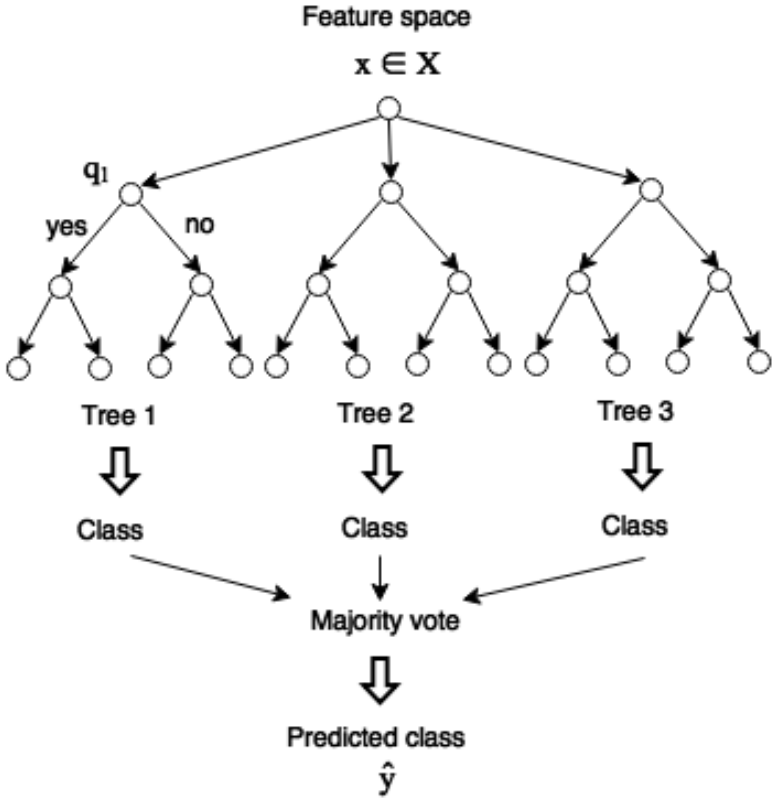

Figure 2.1: Random forest with three trees

By allowing the algorithm to only consider a subset of features when searching for a split-point the overall correlation between the trees is reduced. If the algorithm were allowed to consider all features it would most likely result in a forest where each tree

have similarities and only the most prominent features are used. When performing classification of a point it will be evaluated against the questions of each tree, and the result will be averaged using a majority vote. In general the performance can be improved by increasing the number of trees, however for all classification tasks a point will be reached where the performance improvement is insignificant. Increasing the number of trees at that point will only increase the complexity of the forest while the actual performance remains the same [12].

2.5.5 Support Vector Machine

Support vector machines strive to separate classes by finding boundaries described as hyperplanes. In the best case scenario classes can be separated entirely. Considering a classification model with two classes, then the optimal boundary is found by maximizing the distance from the hyperplane to the closest data points, where each side consist of one class. When the classes are overlapping the goal is instead to find the optimal boundary while accepting some error. To allow for some errors the slack variables ξ are introduced. The optimization problem with hyperplane variables denoted by β can be described as

$$\min ||\beta|| \quad \text{subject to} \begin{cases} y_i(x_i^T\beta + \beta_0) \geq 1 - \xi_i, \forall i \\ \xi_i \geq 0, \sum \xi_i \leq C \end{cases} \tag{2.20}$$

The hyperparameter C is called the box constraint and it is used to tune how much error is allowed. The box constraint will add weights to the slack variables. A large box constraint implies a higher cost of misclassified points while a smaller box constraint allows for more errors. The box constraint is usually described by reformulating (2.20) to

$$\min ||\beta||^2 + C\sum_{i=1}^{N} \xi_i \quad \text{subject to} \begin{cases} y_i(x_i^T\beta + \beta_0) \geq 1 - \xi_i, \forall i \\ \xi \geq 0 \end{cases} \tag{2.21}$$

The optimization problem is solved by forming the Lagrange function from (2.21) and minimizing it with respect to β, β_0 and ξ.

SVM can be used with different kernels, where the kernel type controls how the hyperplanes are created. The linear kernel strives to find a separating hyperplane in the orginial feature space while other kernels initially transform the feature space before finding the hyperplane. When the inverse transform is applied, the decision boundary no longer corresponds to a hyperplane. Two types will be considered in this thesis, linear kernel and radial basis function. The linear kernel results in a decision boundary which is linear in the original feature space. The radial basis funcion results in a decision boundary which is described by exponential functions in the original feature space [12].

2.5.6 Artificial Neural Network

Artificial neural networks extracts linear combinations to explain the relation between the input and the output of the network. By using several layers in the

network, nonlinear relations can be estimated as well, which makes neural networks a powerful classification method. A network with a single hidden layer is shown in Figure 2.2, with inputs X and outputs Y. In a simple feed-forward neural network, each node is fully connected with the nodes of the adjacent layers. Each connection is weighted, making the nodes in the hidden layer and the output layer a linear combination of the previous layer. When used for classification, the final layer is usually connected with a final classification function g_k transforming the output to a class.

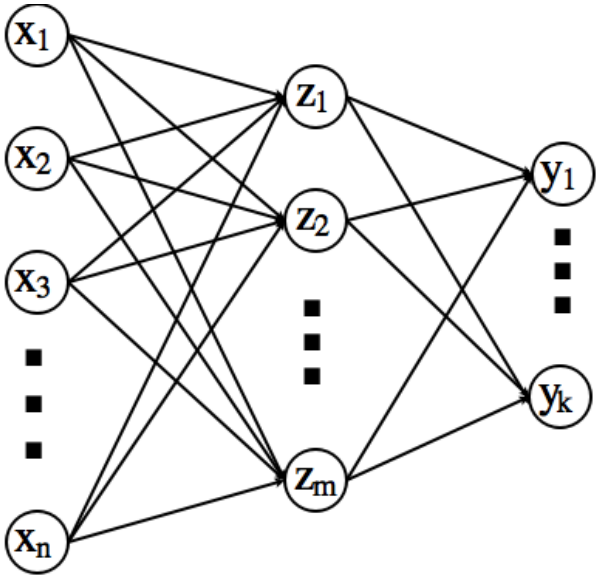

Figure 2.2: Schematics of a single hidden layer, feed-forward neural network.

The linear combinations can be described as follows

$$\begin{aligned} Z_m &= \sigma(\alpha_{0m} + \alpha_m^T X), m = 1, ..., M \\ T_k &= \beta_{0k} + \beta_k^T Z, k = 1, ..., K \\ y_k(X) &= g_k(T), k = 1, ..., K \end{aligned} \tag{2.22}$$

where α_m and β_k denotes the weights connecting the layers and the additional biases α_{0m} and β_{0m}. The activation function σ is usually the sigmoid function, and for classification the softmax function is used for output function g_k.

The network is trained using backpropagation. Backpropagation uses the known schematics of the network to use stochastic gradient descent to minimize the errors of the network. Backpropagation consists of two phases, initially a training input vector is propagated forward through the network. The estimated output is compared with the known output and errors are calculated. These errors are propagated backwards to evaluate the errors of each layer. In the next phase, both sets of errors are then used with gradient descent to update the weights of the network.

One significant problem with training a neural network is the concept of overfitting. If the training process only strives to minimize the training error, then the network will learn the training data set extremely well while any other input could result in low performance. A simple yet efficient approach to avoid overfitting is to continuously evaluate the performance of the network using a validation set. The validation set is a subset of the training set which is not actually used for the training process itself, it is only used to monitor the training process. As long as the performance of the validation set is improved, the training phase should continue. When the validation error is at a minima the training phase should stop [12].

2.6 Evaluation

The performance of each classification method should be evaluated. Cross-validation is performed to evaluate both binary classifiers and the final post-processed result.

2.6.1 Cross-validation

A commonly used evaluation method is cross-validation, it is an iterative procedure used for estimating prediction error. Cross-validation is performed by initially dividing the data into k folds. A classifier is then trained on k-1 folds and tested on the remaining fold. This is repeated until all folds have been used as the test fold and the accuracy is averaged over all folds to estimate the mean and standard error. Cross-validation can be performed in several ways where the main difference is how the folds are chosen.

2.6.2 Performance

The performance of the classification process is evaluated in several steps. Initially the binary classifier performance is evaluated for each method, the measures accuracy, specificity and sensitivity is calculated. The results from the binary classifier is post-processed and the measures true positives, false negatives and false positives are extracted.

2.6.2.1 Binary classifier performance

The binary classifier performance is evaluated for each second. Specificity denotes the amount of accurately classified non-seizure seconds and sensitivity denotes the amount of accurately classified seizure seconds.

$$\text{Sensitivity} = \frac{\text{Amount of accurate seizure classified seconds}}{\text{Total amount of actual seizure seconds}} \tag{2.23}$$

$$\text{Specificity} = \frac{\text{Amount of accurate non-seizure classified seconds}}{\text{Total amount of actual non-seizure seconds}} \tag{2.24}$$

Accuracy is the combination of sensitivity and specificity to give an overall performance measurement. It can serve as an indicator on how important each class is for the total performance. In this thesis where the used data set is very imbalanced it is expected that the total accuracy will be similar to the specificity since the non-seizure movements dominate the data set. For instance, if the specificity is 80% while the sensitivity is 50%, the total amount of misclassified non-seizure seconds will be much larger than the amount of misclassified seizure seconds. This implies that there might be a high amount of false seizure classifications which must be dealt with.

$$\text{Accuracy} = \frac{\text{Amount of accurate classified seconds}}{\text{Total amount of seconds}} \tag{2.25}$$

2.6.2.2 Classification performance

The output of the binary classifiers should be processed before the final classification performance is evaluated. Accurate classification of each second is not the ultimate goal, the task is to identify time segments where seizures occur. The performance is measured in true positives, false negatives and false positives. True positives, TP, denotes how many times the classifier have accurately identified a seizure and false negatives, FN, is the amount of missed seizures. False positives, FP, is the amount of misclassified normal movements, i.e the amount of false alarms. The amount of true positives and false negatives should equal the total amount of seizures in the test set while the amount of false positives is only limited by the amount of measured movements. True negatives are not considered since the concept of normal movements is not defined, normal movements can have varying durations unlike seizures which have a limited duration.

$$TP = \text{Amount of accurate classified seizures} \tag{2.26}$$

$$FN = \text{Amount of missed seizures} \tag{2.27}$$

$$FP = \text{Amount false alarms} \tag{2.28}$$

3
Methods

3.1 Algorithm structure

The structure of the data processing from raw data to final classification is described in Figure 3.1. Initially the accelerometer data is pre-processed to retrieve valid data. The most significant part of the pre-processing is labeling seizures and determining which measured seconds are movements and which seconds are non-movement. All movement seconds are used to calculate features. In the next step the data set is reduced further by only selecting movements from longer periods of time with low activity, which is an approximation to nocturnal seizure detection. Due to the limited data set, all seizures are kept regardless of activity. Depending on evaluation method, either the general leave-one-patient-out or individual leave-one-seizure-out will be evaluated for all classification methods. The result is post-processed to receive classified segments instead of classified seconds since the main task is to identify seizure occurences, not duration and exact onset. The amount of true positives, false positives and false negatives are extracted in the final step.

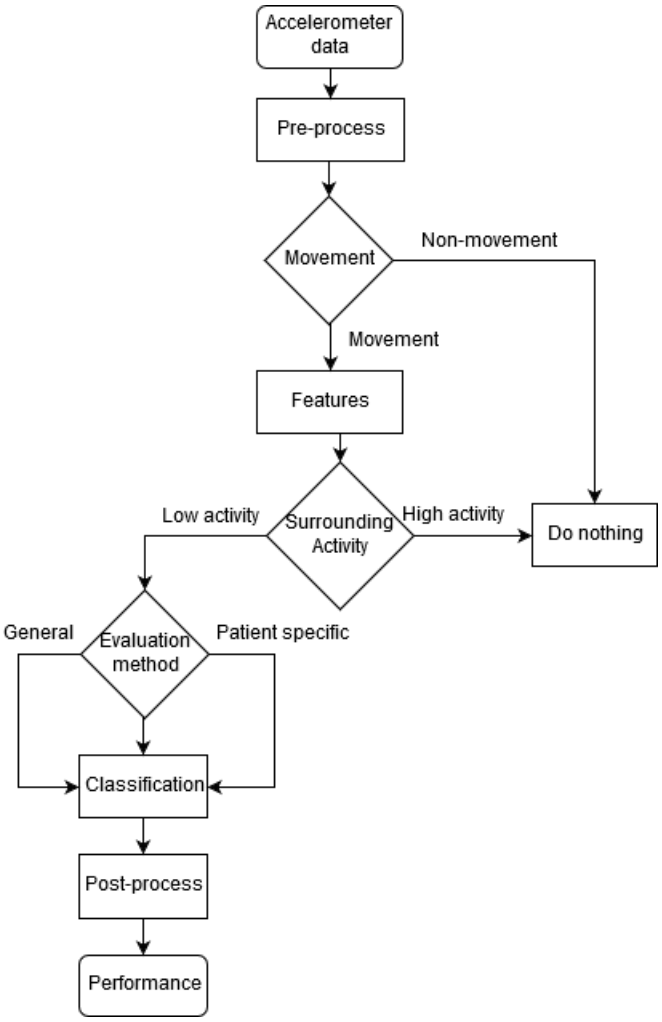

Figure 3.1: Flowchart describing the most significant parts of the classification process.

3.2 Pre-processing

Several days of continuously gathered measurements quickly results in a large data set. One of the most significant steps in the pre-processing is thus to reduce the amount of data. To that purpose the data is initially evaluated to identify segments of motion data and the rest should be discarded. The quadratic mean of the accelerometer data is calculated and evaluated using a sliding window of two seconds with one second overlap. A segment is considered a movement if the standard deviation exceeds a certain threshold. Each measured second is given a movement status which indicates whether it belongs to a movement segment or not. To find the threshold which separates movement from non-movement each segment corresponding to a seizure is evaluated and the standard deviation is calculated. The threshold should be as large as possible to result in a significant data reduction while it should be small enough to ensure that no seizure is removed during the pre-processing.

The validity of each movement segment is evaluated. In some segments there might be interpolation due to data loss caused by sensors losing contact or other malfunctions. As these segments does not provide accurate measurements they are removed. The interpolation is determined by examining the derivative of the accelerometer data, if it is constant then the signal is interpolated. Lastly the motion data is filtered using a Butterworth filter with passband 0.3-22 Hz to remove the gravitational component and high frequences.

3.2.1 Time segments

After the pre-processing the time is piecewise continuous due to the removal of data. For the remainder of the thesis, time segments or segments corresponds to sequences with continuous time. Sequences which are less than ten seconds apart are clustered and considered to be part of the same segment.

3.3 Data reduction

The pre-processing reduces the original dataset by removing non-movement data. However if the data have been gathered for several days one can assume that the amount of seizures is significantly smaller than the amount of regular movements. This implies a very imbalanced dataset which can lead to poorly trained classifiers since many classification methods are trained by minimising the total error. In the worst case the classifier will not learn how to separate the classes at all, instead it assigns the most common class to everything and accepts the small error caused by the few data points belonging to the other class. Balancing the dataset and removing or reducing the bias towards learning a particular class can thus be important to achieve a classifier which gives a reasonable sensitivity.

3.3.1 Activity based reduction

As described in Section 1.3.2 the hypermotoric seizures commonly occurs when the patient is asleep. The initial main focus can thus sufficiently be to focus on low-activity hours. Additionally there is less regular movement than when the patient is awake, which implies that the data set can be heavily reduced.

To that end each patients activity is evaluated. Two cases are investigated. In both cases, due to the limited number of seizures, all seizures are kept regardless of the activity of the segment where it occurs. The activity is calculated using the movement status assigned in the pre-processing of the data.

3.3.1.1 Hour activity

The activity a_i is defined as the amount of movement of each hour i. The activity is calculated using the movement status $m_j = \{0, 1\}$ for each measured second j.

$$a_i = \frac{1}{3600} \sum_{j=(i-1)*3600+1}^{3600*i} m_j \tag{3.1}$$

Hours with larger activity than the threshold 0.1 is discarded. This means that it is acceptable to move for 6 minutes each hour.

3.3.1.2 Quarter activity

The activity is defined as the amount of movement of each 15 minutes.

$$a_i = \frac{1}{900} \sum_{j=(i-1)*900+1}^{900*i} m_j \tag{3.2}$$

If a quarter activity is less than the threshold 0.1 it is considered low activity. If the activity exceeds the threshold the surroundings of the period is investigated. Three quarters, i.e 45 minutes, is evaluated before and after the current quarter. If at least three of those periods are of low activity the current quarter is considered low activity else it is discarded.

$$a_i < 0.1 || (\sum_{i-3}^{i-1} (a_i < 0.1) + \sum_{i+1}^{i+3} (a_i < 0.1)) \geq 3 \tag{3.3}$$

Each quarter activity a_i is evaluated according to (3.3), if the statement is true the quarter is of low activity.

3.3.2 Kernel density estimation

Kernel density estimation is used to model non-seizure movements. Initial evaluation of the feature space, Section 4.3, shows a significant overlap between seizure movements and non-seizure movements. It implies that it is not feasible to expect a decision boundary which separates the two classes entirely. Instead, the method

is used to find a boundary which determines if a time segment consists of only non-seizure movements, however it is not expected to be accurate enough to determine whether an outlier corresponds to a seizure or not.

The decision boundary can optionally be used as a part of the classification process. If used, the data points within the boundary can be classified as non-seizure movements. The classified non-seizure seconds will not be used in further classification methods, neither in the training nor the testing process. When applied it is important to note that the method might accidently classify seizure data of the test set as normal data and the already classified data will not be considered further. However, seizure data points are only misclassified if they are dissimilar to the seizures of the training set. These seizures might arguably be too subtle or too different from the training set to be accurately identified regardless of method, and if they are not removed in this step they might have a negative impact on the remaining classification methods used. The hope is to reduce the amount of false positives in the end, at a hopefully low amount of initial misclassifications.

Using MATLAB [11] and the function `ksdensity`, the probability function can be estimated by evaluating gaussian kernel functions at points covering the range of the data. The estimated density function will be bivariate and estimated at 900 points. The features are individually averaged over channels and over time segments, where time segments denotes episodes with continuous time. The probability density function is estimated for each available feature pair and a boundary is found which covers the largest part of the normal training data while excluding all seizure training data. Initially the seizure boundary is found, that is the boundary which exactly separates seizures from normal movements. To reduce the risk of accidental misclassification of seizure data the seizure boundary is reduced to cover at least 10% less of the normal training data than the original boundary. The decision boundary is applied to the time segments of both the training data and the test data and each second is given a status to indicate if it is likely or less likely that it belongs to a segment which only contains normal movements.

3.3.3 Subtle seizures

Seizures with subtle and short motoric manifestations can be very similiar to regular movements. These seizures can have a large impact on the total accuracy of the classifier. When training a classifier on these subtle seizures the classifier will most likely result in a classifier prone to produce false positives as many regular movements will be similar to said seizures.

These seizures can either be included or removed in the training and evaluation process. If they are included it indicates that the goal is to have a classifier sensitive enough to accurately identify short and subtle seizures. However it should be noted that the task is made more difficult and will most likely reduce the specificity. If instead the seizures are removed from the data sets it is under the assumption that

it is acceptable to never notice these short and subtle seizures and the focus is to have a high sensitivity for severe seizures and a higher specificity. In this thesis, all seizures have been included regardless of their subtlety.

3.3.4 Balancing dataset

A fully balanced training set with 50% normal movements and 50% seizure movements implies that there is no preference, it is of equal importance to accurately classify each data point regardless of which class it belongs to. The test set should not be balanced since the timeline will not be kept intact and it will make the result difficult to interpret. Additionally, when the classifiers are used in a real-life application it will never be presented balanced data.

Two methods are used to balance the training data set. In the first method the dataset is balanced by random selection of normal data until there is an equal amount of normal data and seizure data. In the second method the normal data is averaged using K-means clustering. The euclidian distances in the normal data set are evaluated to create K clusters, where K is the amount of seizure data points. The normal data is then exchanged by the centroid locations.

3.4 Features

The pre-processed data is used to calculate features as described in Section 2.2. Most features will be represented by two vectors where each vector corresponds to one sensor, having values for each second of the pre-processed data set. Correlation only consist of one feature vector since the calculations consider both sensors. Additionally, the energy content is evaluated using 14 frequency bands resulting in 28 feature vectors. For convenience, the frequency bands are clustered into four separate features, where each feature consists of several frequency bands. By representing the energy content using multiple vectors rather than one representing the entire band, it enables the possibility to weigh or favor the bands differently. The frequency bands range from 0.75 to 11.25 Hz.

3.4.1 Feature selection

The feature selection methods are used to select a pre-defined number of features. The evaluated number of features are 2, 5, 10 and the original case with no feature space reduction.

3.4.1.1 Forward Selection

The optimal set of features is model dependent. A feature which is important for one classification method might be irrelevant for another. In the GTCS thesis an optimal feature space was generated for each method, and the result clearly shows

differences among the classifiers. For instance, KNN favored SMA while logistic regression and random forest did not [9]. Given their result it is of importance to consider the classification algorithm when using forward selection. To that end, the forward selection algorithm is implemented to use the intended classification algorithm. If the reduced feature space should be used with a KNN classification model, the same setup will be used in the process to find which set of features should be used.

3.4.1.2 Overlap Integral

The overlap integral method is used to find the features with smallest overlap. The MATLAB [11] function `histogram` is used to estimate the probability density functions for each feature f using N equally spaced bins. The overlap is calculated as the sum of the products of the normal bins n_i and the seizure bins s_i. A small overlap indicates a well-separated feature. For this thesis, N = 50 is used to calculate the histograms.

$$overlap(f) = \sum_{i=1}^{N} n_i s_i \tag{3.4}$$

To reduce redundancy of the overlap integral reduction, the method is modified to consider pre-defined similarities, the method can only choose one feature among similar features. The similarities considered is described by the list below, the method can only choose one feature from each row.

- Spectral Edge 80, 90, 95
- Mean value, RMS
- Variance, Standard deviation

3.5 Classification

The reduced and pre-processed accelerometer data and the extracted features are used to train several classifiers. The original case with all features is considered as well as different combinations of feature selection methods and kernel density estimation.

Artificial neural networks is used to evaluate pre-processed accelerometer data instead of the calculated features. Perhaps there are some form of pattern and information in the original data which is not entirely covered by the feature set. Using neural network on raw data is an attempt to learn seizure characteristics directly instead of relying on designed features. Naturally, this approach is not combined with feature selection methods and kernel density estimation as they both require the calculated features. The binary result of the neural networks is processed using the same post-processing layer as the feature classifiers to give comparable results. This includes the veto which are feature based. Hence, the total performance of the neural networks depends on both accelerometer data and features.

In the previously mentioned classification algorithms, each data point represent one second of data. In the neural network approach each sample is considered instead, an input equivalent to one second will consist of 50 samples. The neural network will be trained on low activity data which have not been further balanced, which makes it important to evaluate whether the network learns to recognize seizures or not. By assigning error weights to the training data, the class with lesser occurence can be favored during the training process. The error weights basically control how much the network should penalize an error, by assigning seizures a larger weight it will have a larger impact on the training process than normal movements. The training process is performed by iteratively changing the error weights in favor of seizures until the network have learned to recognize at least 50% of the seizure data in the validation set.

3.5.1 Binary Classification

A set of parameters are evaluated for each classification algorithm. Additonally, each parameter is evaluated in combination with forward selection, denoted FS, overlap integral reduction, OI, and kernel density estimation, KDE.

3.5.1.1 Classification algorithms

The parameters and combinations to be evaluated for each classification algorithm is shown in Table 3.1.

Table 3.1: Combination table.

Primary method	Parameter 1	Parameter 2	KDE	FS	OI
KNN	Neighbors 2, 5, 10				
			✗		
				✗	
			✗	✗	
					✗
			✗		✗
Random Forest	Trees 10, 30, 50				
			✗		
				✗	
			✗	✗	
					✗
			✗		✗
Logistic regression					
			✗		
				✗	
			✗	✗	
					✗
			✗		✗

SVM	Kernel linear, rbf	Box constraint 1, 100	✗ ✗ ✗	✗ ✗	✗ ✗
ANN	Hidden neurons 2, 5, 10, 18	Hidden depth 0, 1, 2			

3.5.2 Post-processing

In order to evaluate the perfomance as described in Section 2.6.2, a post-processing structure is developed. The result from the binary classifiers should be processed further in order to be able to better interpret the result. The binary classifier results in a classification for each tested second, but since the seizures are longer than this it is not desirable to directly present the result for each second. Several misclassifications within a short time period should not result in several false positives and several true seizure classified seconds should not neccessarily be equivalent to the same amount of true positives. Additionally, to reduce the amount of false positives two veto filters are evaluated. The post-processing structure is described in Figure 3.2.

The post-processing structure is created manually given knowledge about the performance and ideas of how to improve it. Given a more complete data set another learning approach could have been evaluated. It is possible that this structure could have been modeled using a hidden Markov model instead. A hidden Markov model consists of a set of states and transitions, much like the flowchart, where there are some observable outputs available but the states themselves are hidden [14].

3.5.2.1 Median filter

Initially the result from the binary classifier is filtered to make the classification smoother, this is an important step towards segment classification rather than one separate classification for each second. Each second is evaluated, if there are at least two seizure markings within a five second window the second should be seizure marked. If not, the second should be marked as non-seizure. In Figure 3.3 the effect of the filter is shown, when there is only one seizure classification within the window it is removed and several seizure classifications have been clustered.

3.5.2.2 Veto

The classification can be further evaluated with the purpose to reduce false positives. Two approaches are investigated using the features correlation, entropy and

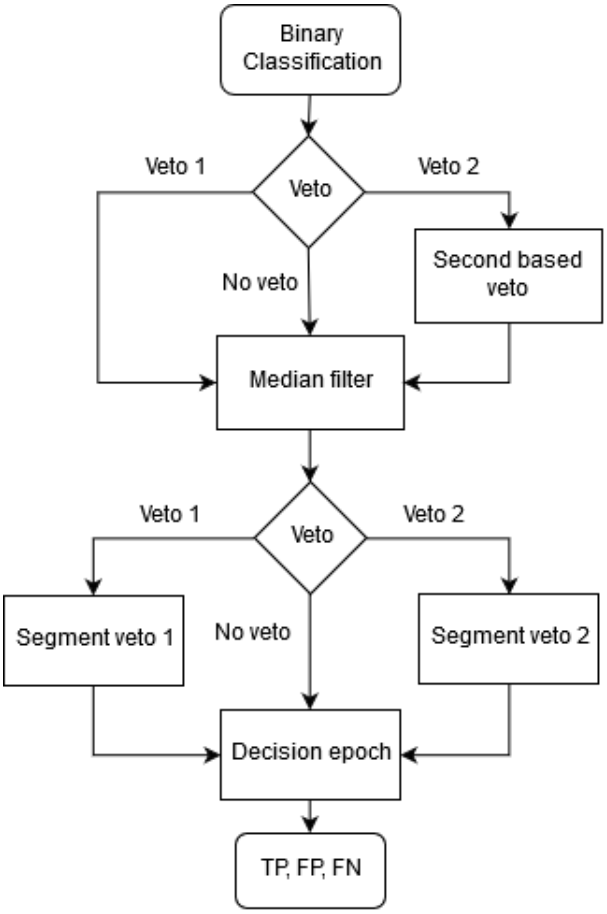

Figure 3.2: Flowchart describing the post-processing from the binary classification result to final classification performance.

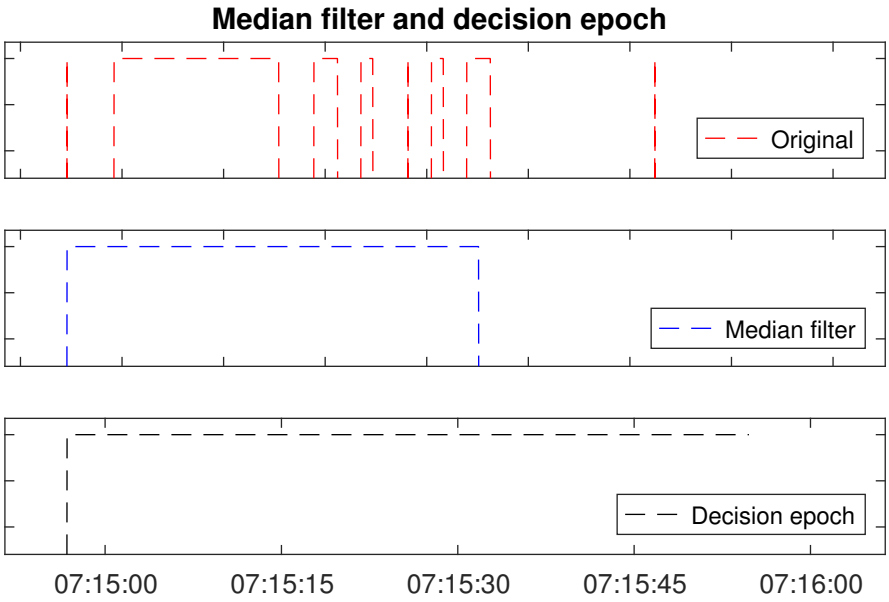

Figure 3.3: The effect of the median filter and decision epoch of a binary classification result. Each second is classified, where a seizure classified second is represented by a peak and normal by a valley.

frequency peak. Initially the seizures of the training set are considered and the mean value of correlation, entropy and frequency peak are calculated, denoted $\bar{c}$, $\bar{e}$ and $\bar{f}$. These means will be used as a veto, if a classified seizure segment does not fulfill the veto requirements the seizure classification will be removed. Two different vetos are evaluated.

Veto 1. Segment based veto Each seizure classified segment is considered individually. The mean values of entropy, correlation and frequency peak are calculated over both sensors and over the segment and compared to the means of the training set.

if $(\text{mean}(\text{correlation}) < 0.6\bar{c} \vee \text{mean}(\text{entropy}) \leq 0.5\bar{e}) \wedge \nexists\ \text{freqPeak} > 0.8\bar{f}$
    class = Normal

Veto 2. Second and segment based veto In addition to a modifed version of Veto 1, each seizure classified second is initially considered individually before the median filter.

if $(\text{correlation} < 0.8\bar{c} \vee \text{entropy} \leq 0.7\bar{e}) \wedge \nexists\ \text{freqPeak} > \bar{f}$
    class = Normal

Afterwards the median filter is applied and each resulting seizure segment is considered as in Veto 1.

if (mean(correlation) $< 0.8\bar{c} \vee$ mean(entropy) $\leq 0.7\bar{e}) \wedge \nexists$ freqPeak $> 0.8\bar{f}$
    class = Normal

3.5.2.3 Decision epoch

As previously mentioned the goal is to reach a segment classification rather than accurate classification of each measured second, estimating exact onset and duration is not considered. As seen in Table A.1 the HMS seizure with longest duration is 32 seconds, however according to literature it is not uncommon for a seizure to last for up to one minute. Additionally, according to the data set there is at least a couple of minutes after a seizure before another seizure occurs. The shortest duration between seizures occurs in patient 37, there is three minutes and 37 seconds between seizure three and four.

The known seizure statistics and literature implies that it is unreasonable to seizure classify several segments in close proximity to each other. Such segments are either due to several misclassifications within a short period of time or the actual seizure is noticable over several segments. Thus the final processing step is to modify each cluster of seizure markings to be of length 90 seconds regardless of the initial amount of markings. Additonally, after each seizure classified cluster it is not allowed to have a seizure marking within 60 seconds. The effect is visualized in Figure 3.3. Note that due to the piecewise continuous timeline the shown decision epoch is shorter than 90 seconds, if there had been additional valid segments within the 90 second window it would have been marked as well.

3.5.3 Performance

The post-processed result is compared to the true classes of the data set. If a seizure classified segment overlaps an actual seizure segment it results in one true positive. If a seizure classified segment have no overlap with an actual seizure it is a false positive. If an actual seizure segment does not have any overlapping seizure classified segment it results in one false negative.

3.6 Cross-validation

The methods are evaluated using cross-validation and two different approaches will be investigated, the general leave-one-patient-out approach and patient specific leave-one-seizure-out. Leave-one patient-out denotes the case where each fold consists of one patient, this is a general method where each patient will be evaluated using classifiers trained on the remaining patients. In this thesis continuous time is of importance since movements and seizures have duration and thus it is not desirable to choose the folds randomly. The timeline is automatically kept intact using leave-one patient-out. The patient specific classification is performed using one patient at a time by creating a fold for each seizure. The normal data is divided according to segments with continuous time and randomly distributed so that each

fold has roughly the same amount.

Patient 36 and 39 have three and two seizures respectively, the corresponding patient specific classification will be trained on a very small amount of seizure data. Leave-one-seizure-out will be performed, but it is important to note that the result might not be very general given the small amount of training data.

Furthermore, as described in Section 2.5.6, artificial neural network requires one additional fold for validation to reduce overfitting. When leave-one-seizure-out is performed one fold is used for testing, and the remaining data is divided into a validation set and a training set. The training set will consist of 85% of the training data and the remaining 15% is used for validation. Due to the imbalanced set, precautions are made to ensure that there is seizure data in both sets. The seizure data and the normal data is divided separately instead of dividing the entire data set at random. When evaluating the general leave-one-patient-out the same method is used to divide the training data into a training set and validation set.

4
Results

4.1 Seizure statistics

The HMS statistics aquired during the pre-processing is presented in Table A.1. Seizure time represents the duration of the motoric manifestations. The maximum standard deviation indicates which threshold is suitable for separating movement from non-movement. If the chosen threshold is larger than the standard deviation of a seizure then it will be discarded during the pre-processing of the data. The maximum threshold allowed is 0.076 g; in order for all seizures to survive the pre-processing with some margin, the threshold for HMS pre-processing is set to equal 0.06 g.

For comparison, when processing the GTCS data in the same manner the maximum threshold allowed equals 0.62 g, which is ten times larger than the chosen threshold for HMS. The statistics for the GTCS data set is shown in Table G.1.

4.2 Low-activity build

The processed data is reduced according to Section 3.3.1, it results in a distribution of seizure seconds according to Tables 4.1 and 4.2. The column "LA seizure" corresponds to the amount of seizure seconds in low activity segments while "HA seizures" corresponds to the amount in high activity segments. There are only low activity seizures available from patient 37 when hour basis is used. The higher resolution of the quarter basis results in low activity seizures from all HMS patients but patient 39. Hence, the hour build is not considered valid and only the quarter based activity will be evaluated. However as previously mentioned, all seizure data will be kept regardless of the activity of the corresponding segment. Additionally, the column "Duration" corresponds to the duration of the low activity build.

Table 4.1: Low-activity build statistics, hour basis.

Patient	Duration [h]	LA seizure [s]	HA seizure [s]
21	16	0	85
36	24	0	83
37	49	217	123
39	14	0	31
total	103	217	322

Table 4.2: Low-activity build statistics, quarter basis.

Patient	Duration [h]	LA seizure [s]	HA seizure [s]
21	26.50	49	45
36	33.25	19	64
37	50.00	229	99
39	17.50	0	31
total	127.25	297	239

Further data is extracted from the quarter basis low-activity build to show the amount of seizure data compared with the amount of normal data. For each patient, the amount of seizure data is 1.3%, 1.4%, 7.0% and 1.0% respectively. This confirms the claimed fact that the data set is very imbalanced.

4.3 Features

The normal and seizure distributions of each feature described in Section 2.2 is evaluated using histograms. Examples are shown in Figures 4.1 to 4.4, where all quarter based low activity build data is used to estimate the distributions. The features have not been normalized, however they will be before usage for classification. There is a clearly visible difference between the distributions but there is also a significant overlap for all features. The difference indicates that it is possible to separate the classes to some extent, but due to the overlap it will be difficult to do so with perfect accuracy.

Figure 4.1 confirms the expected behavior of the feature correlation, a seizure commonly results in a motor behavior in both arms. According to the distributions, if the correlation of a segment is low it is most likely a non-seizure movement, while a higher correlation indicates that the segment might contain a seizure. Similarly, entropy indicates high activity in seizures while there is a lot of variation in the normal non-seizure movements. This was expected, normal movements have a wide range of intensity and activity, some movements are subtle while others are not. The high activity level of seizures is verified by the feature Spectral Edge Frequency, Figure 4.2, which clearly indicates that seizures contain frequencies from a larger spectrum than normal movements.

The feature frequency peak, Figure 4.3, indicates that if a segment have any large frequency peaks it is likely to be a seizure segment. The newly added feature Jerk, Figure 4.2, confirms the hypotheses that the behavior of a seizure is more erratic than normal movements, and as a result generally higher jerk. Vector magnitude, mean value and standard deviation results in similar distributions where normal movements are focused around a low feature value while higher values increases the

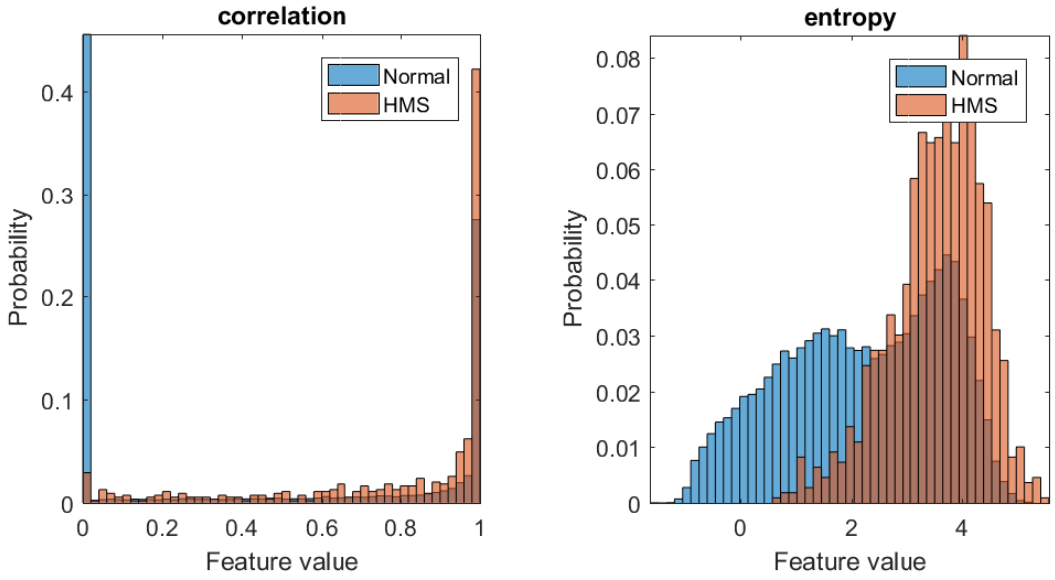

Figure 4.1: Left: Correlation of all HMS patients. Right: Entropy of all HMS patients.

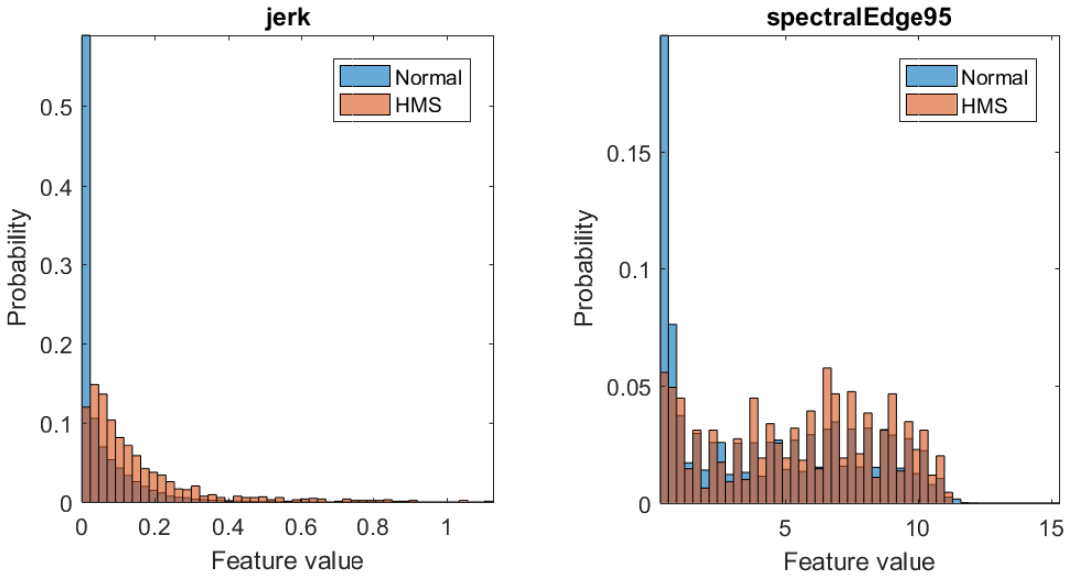

Figure 4.2: Left: Jerk of all HMS patients. Right: Spectral edge frequency 95% of all HMS patients.

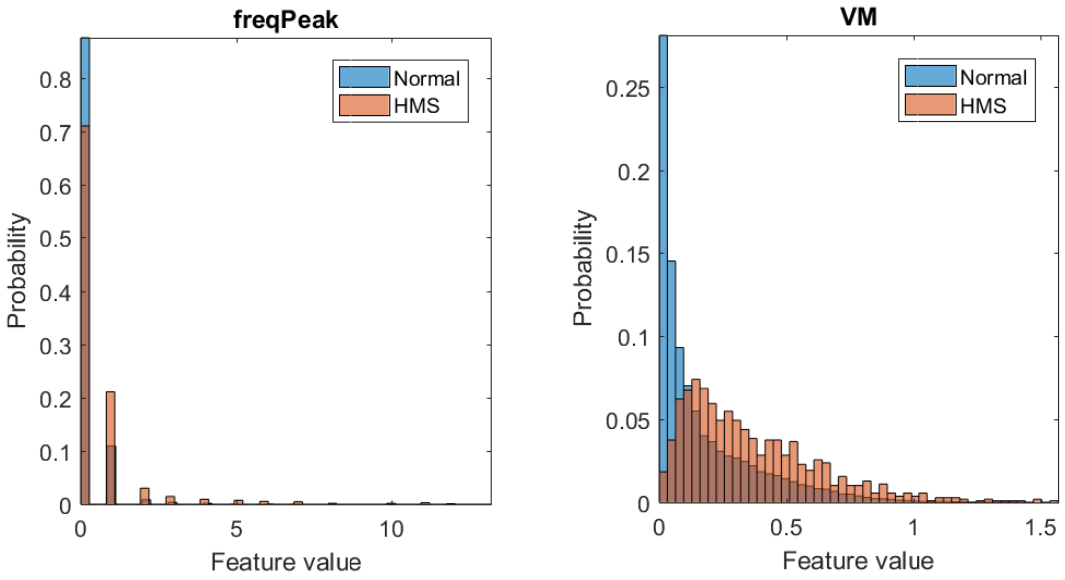

Figure 4.3: Left: Frequency peak of all HMS patients. Right: Vector magnitude of all HMS patients.

probability of a seizure.

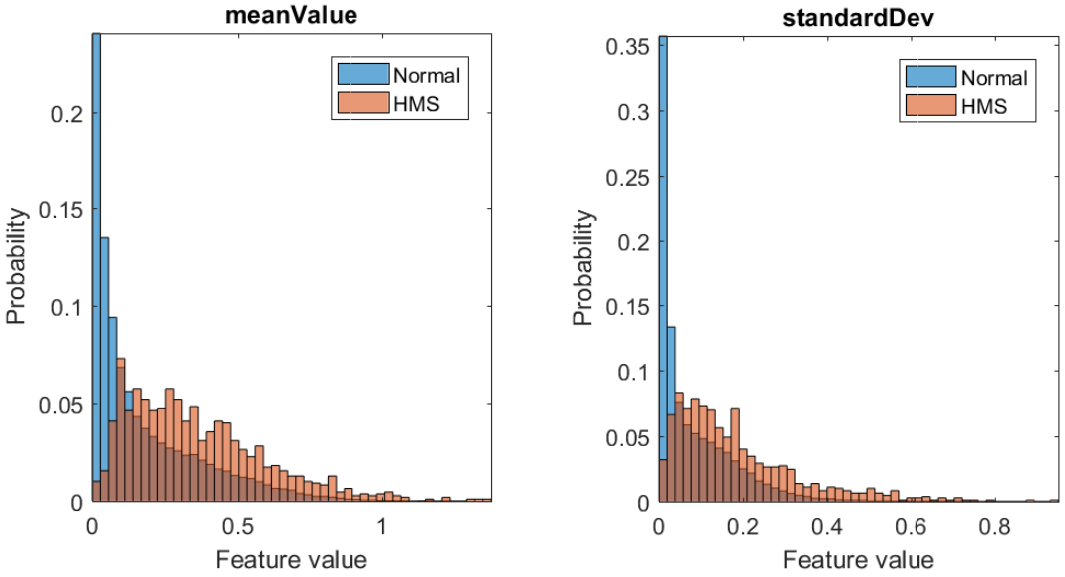

Figure 4.4: Left: Mean value of all HMS patients. Right: Standard deviation of all HMS patients.

4.4 Subtle and atypical seizures

Subjective evaluation of accelerometer data, video and the post-processed result indicated that some of the seizures are subtle or atypical. Some of these seizures are visually very similar to normal movements but given the activity in the video-EEG they are indeed involuntary movements caused by epilepsy, however very subtle. Hence, some seizures are considered subtle due to short duration and a low amount of activity in the accelerometer data, the seizures seems to consist of only a few movements. The seizures which are considered subtle or atypical are seizure 3 and 7 for patient 21 and seizures 6, 15, 16, 17 and 18 for patient 37. To illustrate the differences, a few examples are shown. First, Figure 4.5 illustrates the difference by plotting the histograms of the feature entropy. The histogram to the left consists of all typical seizures from all patients while the histogram to the right consist of the atypical seizures. The seizure distribution of the atypical seizures is shifted to the left, indicating a generally lower activity than the case with only typical seizures.

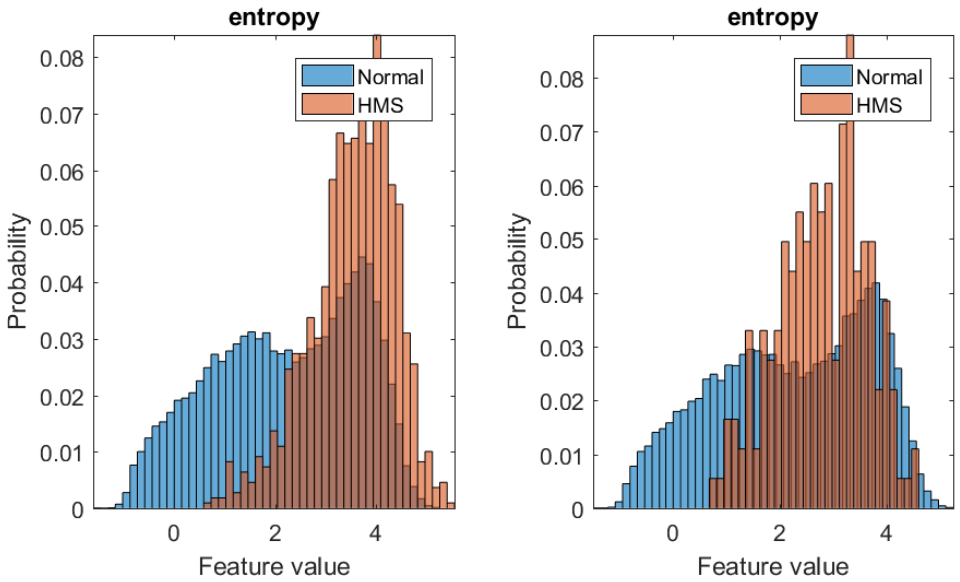

Figure 4.5: Left: Entropy for all typical seizures. Right: Entropy for all subtle and atypical seizures.

A typical seizure is shown in Figure 4.6, even though the duration is short there is a lot of activity in both sensors. The same patient have several seizures with a subtle behavior, an example is provided in Figure 4.7 where there is much less activity, the patient is barely moving at all. Figure 4.8 illustrates a subtle seizure of patient 21 in proximity to normal non-seizure movements. There is no obvious difference, the seizure is very similar to the normal movements. Additonally, patient 21 have a seizure where the signal is interpolated in the middle of the seizure, Figure 4.9.

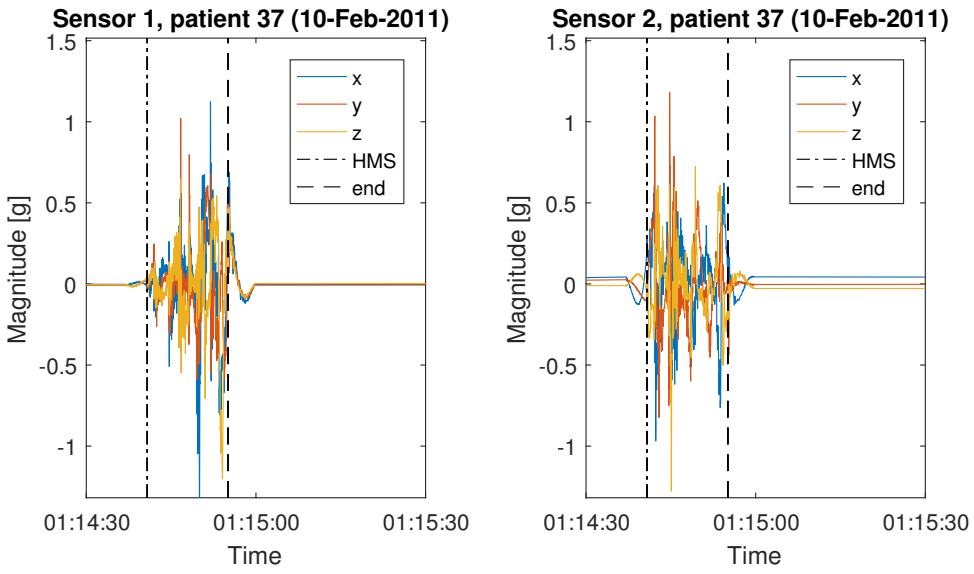

Figure 4.6: The motor behavior of the fifth pre-processed hypermotor seizure of patient 37. Typical behavior.

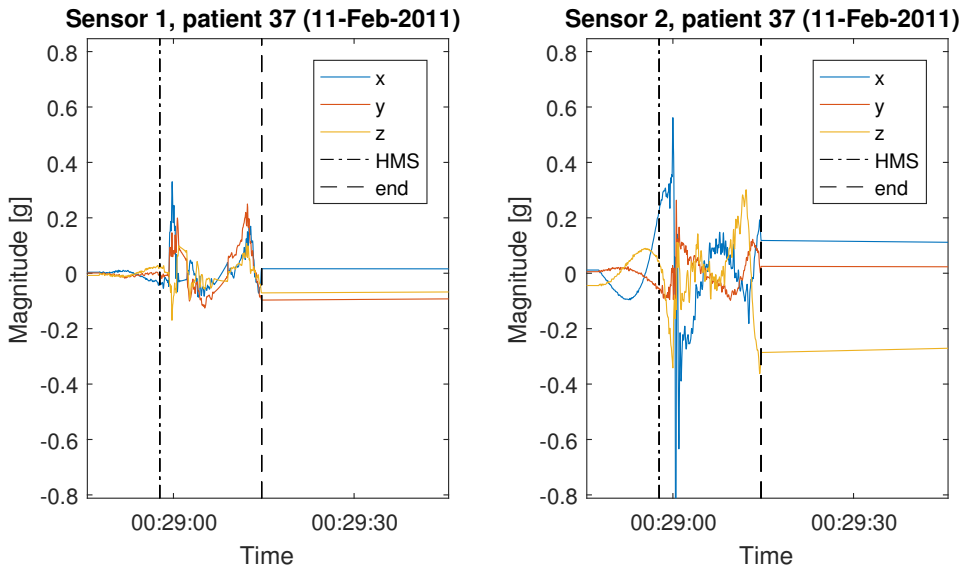

Figure 4.7: The motor behavior of the sixteenth pre-processed hypermotor seizure of patient 37. The seizure is very subtle.

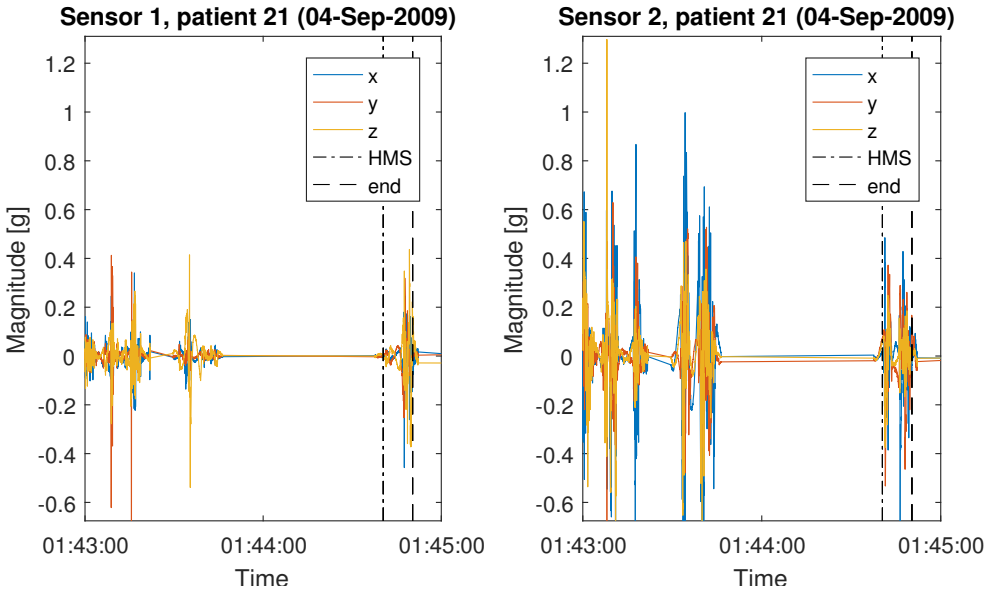

Figure 4.8: The motor behavior of the third pre-processed hypermotor seizure of patient 21. The seizure is similar to normal movements.

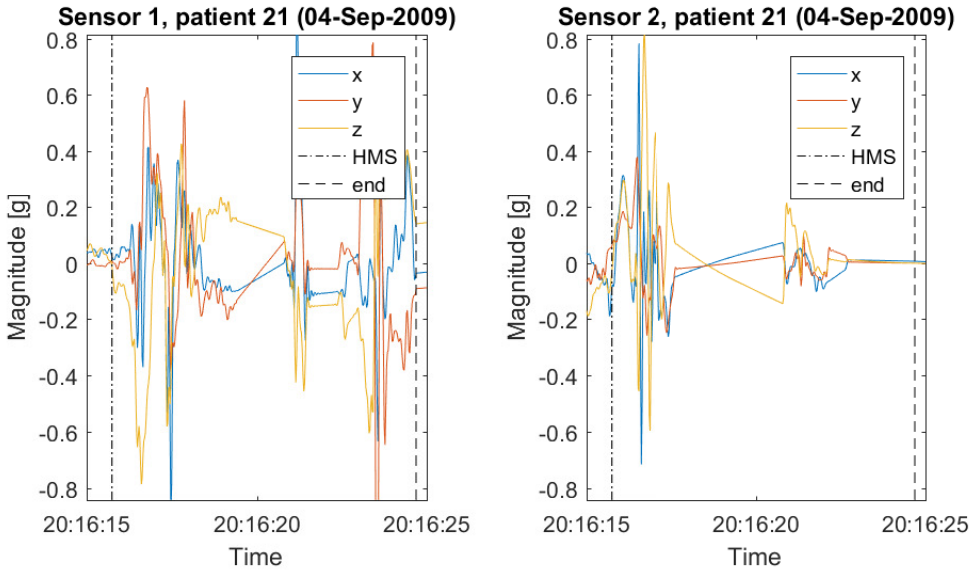

Figure 4.9: The motor behavior of the seventh pre-processed hypermotor seizure of patient 21. The signal is interpolated in some parts.

4.5 Kernel Density Estimate

The KDE method described in Section 3.3.2 is evaluated in both the general case and the patient specific case. From each training set all combinations of feature pairs are evaluated to find the decision boundary which provides the best separation of normal segments and seizure segments. Each patient is considered separatly

for evaluation of the performance, the optimal decision boundaries are applied to the corresponding patient data sets and the amount of normal data and seizure data within each boundary is found. Whether the data belongs to the training data or testing data is not differentiated in the evaluation here, but when the method is used in the classification process the data within the boundary from the training set will not be used for further training and data from the test set will be classified according to whether the data is inside the boundary or not. Additionally for evaluation purposes only, if the same feature pair is used multiple times the result is averaged over the corresponding decision boundaries. Furthermore, the total average over all cross-validation folds and corresponding decision boundaries is calculated.

The performance is summarized in tables and figures. The tables show the amount of normal data and seizure data within each decision boundary chosen by the algorithm. The figures illustrates the performance by plotting the seizure boundary for each fold using a red dotted line and the average boundary in black. The average boundary is not used during the classification, it only strives to illustrate the average behavior of the method. Blue dots represents the normal segments and red dots represent seizure segments. For visual evaluation of the performance it is the segments in close proximity to the boundaries which is of interest since the shape is limited by the seizures that are close to the dominating cluster of normal segments. To display the boundaries with higher resolution, all outlier segments are not shown.

4.5.1 General performance

The general performance for each patient is shown in Table 4.3. Given the four folds, this could have resulted in four different feature pairs. However, variance and correlation was optimal for three folds and the decision boundary for the remaining fold was optimal given the feature pair correlation and frequency band 1 (low frequencies). For visual evaluation, plots corresponding to the feature pair variance and correlation is shown in Figure 4.10 and 4.11.

Table 4.3: General KDE performance

Patient		Feature 1	Feature 2	Normal data	Seizure data
		Correlation	Frequency band 1	27.6%	19.1%
21		Variance	Correlation	40%	0%
	Average:			36.9%	4.79%
		Correlation	Frequency band 1	20%	0%
36		Variance	Correlation	23.2%	0%
	Average:			22.4%	0%
		Correlation	Frequency band 1	27.6%	0%
37		Variance	Correlation	27.6%	0%
	Average:			27.6%	0%
		Correlation	Frequency band 1	14.8%	0%
39		Variance	Correlation	13.6%	0%
	Average:			13.9%	0%

The general segments can be evaluated by inspection of Figures 4.10 and 4.11. For patient 36 and 39 the seizure segments are outliers while patient 21 and 37 have seizure segments which are placed in the cluster of normal segments. The performance of the KDE method is highly affected by these seizures, either or both patients are always used in the general training process, hence the decision boundary is reduced to cover only a small portion of the normal data. This explains the low percentage of normal data shown in Table 4.3. For Patient 21 a seizure segment is removed due to the KDE method. The seizure which is removed is the seizure which is partly interpolated and previously shown in Figure 4.9. When this seizure is in the test set, the amount of classified normal segments is increased when compared to the cases where it is not. This indicates that if the seizure had not been considered at all, the performance where it is a part of the training set might be improved.

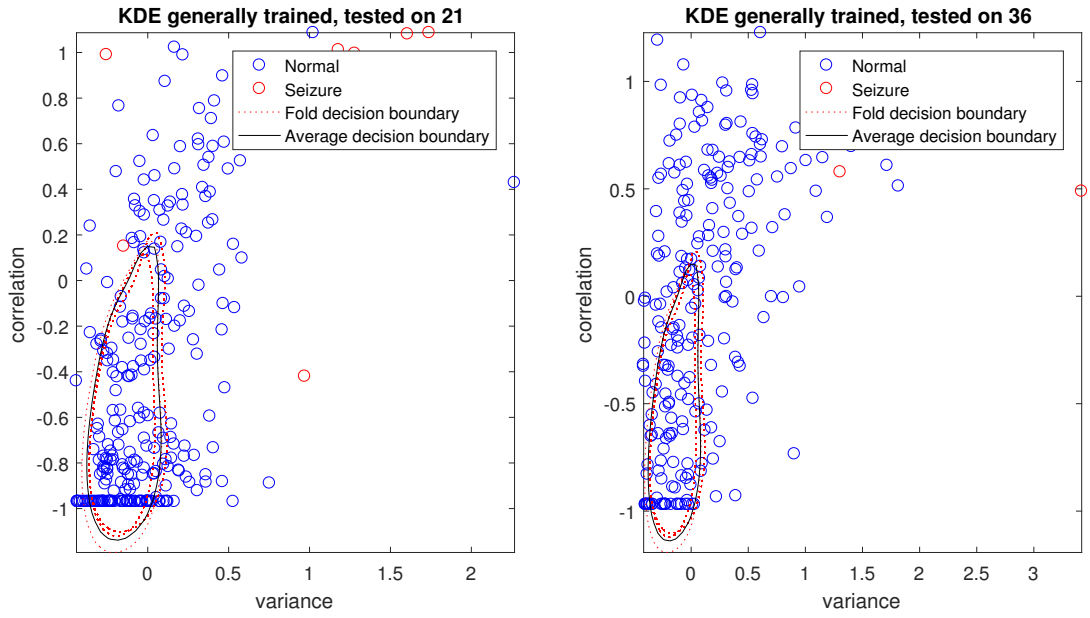

Figure 4.10: General KDE, estimate based on three patients and evaluated using the remaining patient. Left: Evaluated using patient 21. Right: Evaluated using patient 36.

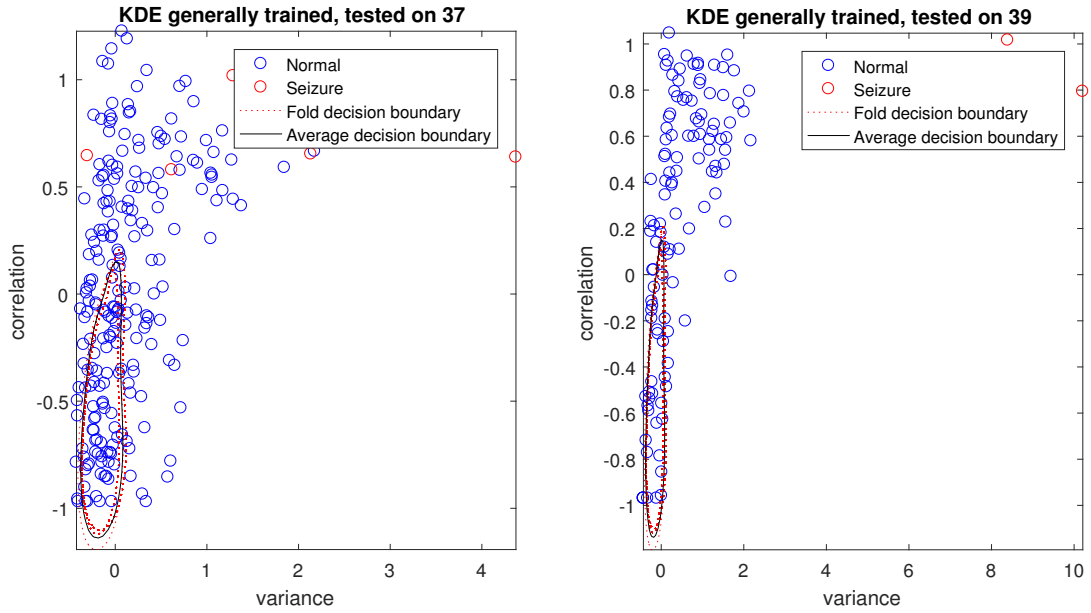

Figure 4.11: General KDE, estimate based on three patients and evaluated using the remaining patient. Left: Evaluated using patient 37. Right: Evaluated using patient 39.

4.5.2 Patient specific performance

The patient specific performance is evaluated as in the general case. The performance is evaluated using the fold structure described in Section 3.6, the method is iteratively trained using all folds but one and tested on the remaining fold. Hence, each patient can result in as many original feature pairs as there are seizures. This is however not always the case, for patient 21 and 37 some pairs are used for several folds. The performance is shown in Table 4.4.

Table 4.4: Patient specific KDE performance

Patient		Feature 1	Feature 2	Normal data	Seizure data
21		Variance	Correlation	62.8%	0%
		RMS	Correlation	64.4%	0%
		Stddev	Jerk	84.2%	19.1%
		Mean Value	Correlation	69.3%	0%
	Average:			67.3%	2.39%
36		VM	Variance	78.1%	0%
		Spectral Edge 80	RMS	86.4%	0%
		Spectral Edge 95	Stddev	81.8%	0%
	Average:			82.1%	0%
37		Correlation	Jerk	33.7%	0%
	Average:			33.7%	0%
39		VM	Stddev	89.5%	0%
		VM	Spectral Edge 90	89%	0%
	Average:			89.3%	0%

Patient 21 results in four feature pairs. By inspection of the left plot in Figure 4.12 it appears that the fold decision boundary is missing. This is not the case, when a specific feature pair is only chosen once the average decision boundary is equvivalent to the fold decision boundary. As in the general approach, when the interpolated seizure is in the test fold it is classified as a normal segment. This can be seen in Figure 4.12 to the left, a seizure segment is clearly placed within the decision boundary. However, the amount of accurately classified normal segments is increased when compared to the cases when it is a part of the training set. According to the right plot in Figure 4.12, the decision boundary is limited by one seizure segment which corresponds to the same seizure, the boundary could have covered more data if the seizure had not been included. As in the general case, this implies yet again that the performance could be improved if it had not been a part of the training set. This clearly indicates that this specific seizure is very dissimilar to the remaining seizures, regardless of training method.

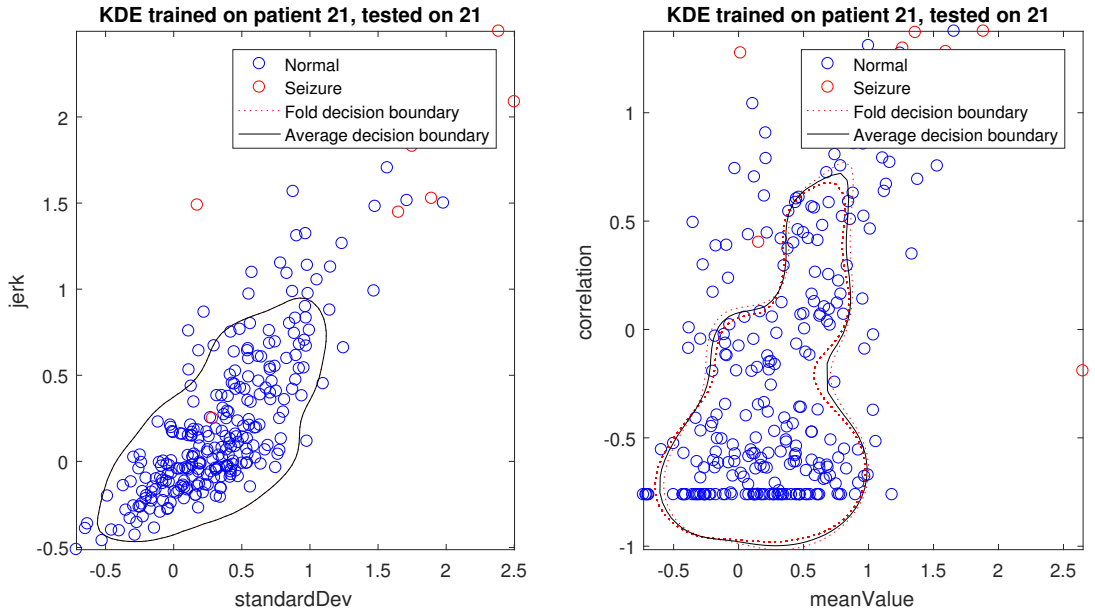

Figure 4.12: Patient specific KDE, estimate and evaluation using patient 21.

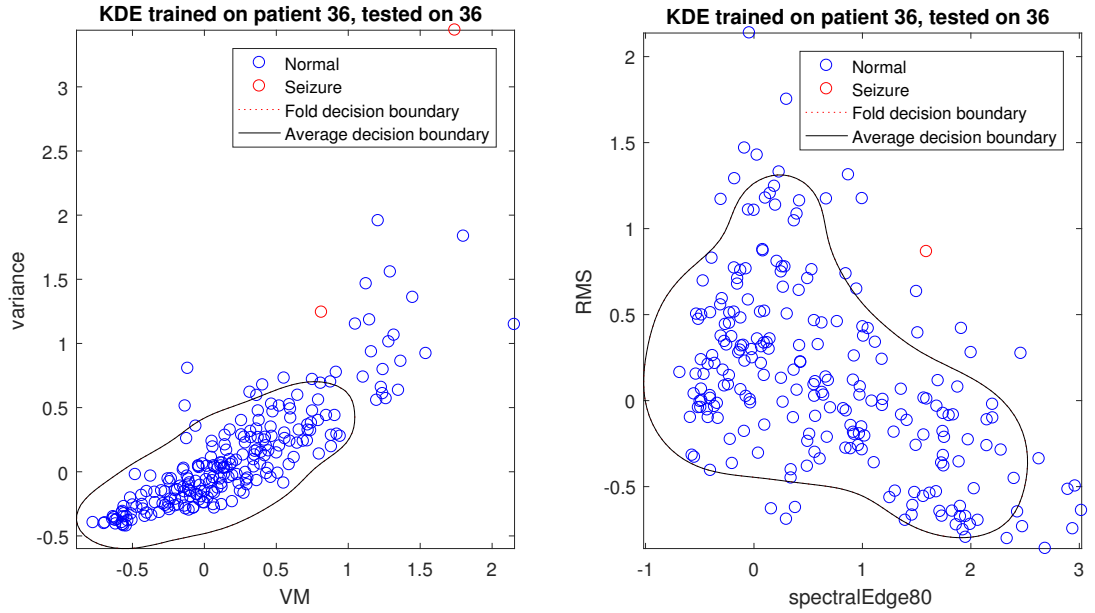

Figure 4.13: Patient specific KDE, estimate and evaluation using patient 36.

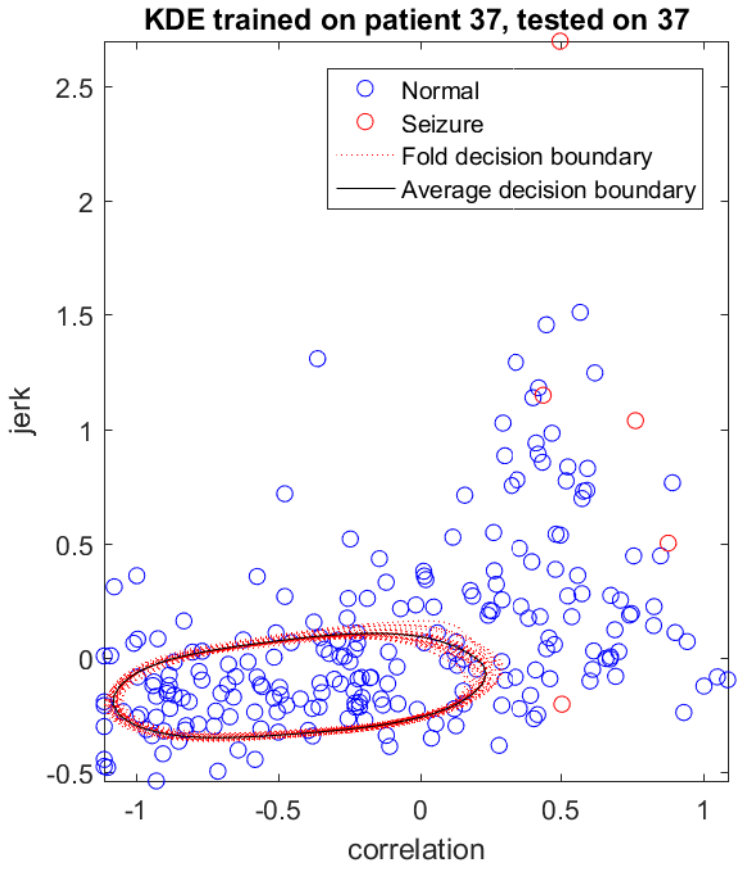

Figure 4.14: Patient specific KDE, estimate and evaluation using patient 37.

One of the seizure segments of patient 36, Figure 4.13, is in close proximity to the non-seizure movements while the remaining two are not. Regardless of fold, no seizure segment is within the boundary. According to Table 4.4, the boundaries cover approximately 80% of the data, making it an efficient method. In contrast, by evaluating Table 4.4, it is evident that patient 37 is not very well suited for the use of KDE. The performance is far worse than for the other patients. Only 34% of the normal data is classified as such while the others reach over 60%. This is due to poor feature separation, there are subtle or atypical seizures which are too similar to non-seizure movements for the method to be efficient.

According to Figure 4.15, by visual inspection patient 39 is the only patient where the seizure segments and normal segments can be separated entirely given the used set of features. If the method had not been implemented to reduce the boundary it would cover all normal segments. Furthermore, perfect separation implies that KDE could be used as the sole classification method for patient 39, without need for further processing. As mentioned in Section 2.5.1, this approach was considered in articles [7] and [8]. This is supported by Table 4.4, patient 39 results in the best performance.

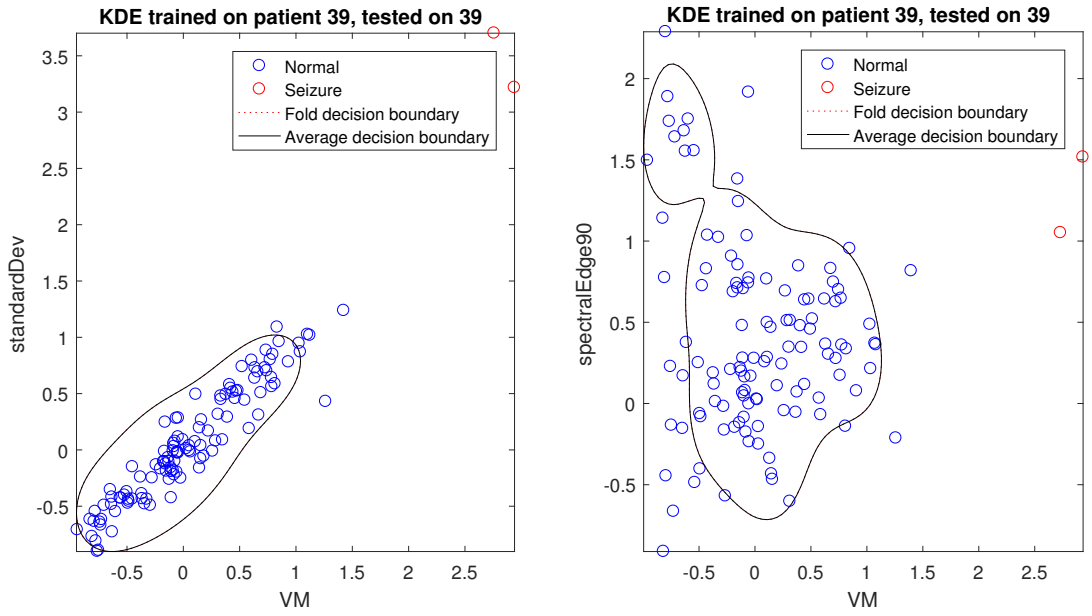

Figure 4.15: Patient specific KDE, estimate and evaluation using patient 39.

4.6 Cross-validation

This section covers the performance aquired using the two cross-validation approaches mentioned, the general leave-one-patient out compared with the patient specific leave-one-seizure-out. The result is shown for each patient, one classification method at a time.

The complete result of the binary classification methods and post-processed result of the general approach can be found in Appendix B and the patient specific approach in Appendix C - F. The following tables have been reduced to only compare KNN with ten neighbors, random forest with 50 trees, SVM with linear kernel and box constraint 1, and artifical neural network with all considered structures. Furthermore, the feature reduction methods will only be shown for the original case and ten features. No combination of KDE and feature reduction is shown since the result is generally quite poor.

4.6.1 Initial evaluation

All combinations described in Table 3.1 leads to a very large set of results. To improve readability and the results have been reduced significantly. All tables have been briefly evaluated to find general performance patterns and a set of combinations are selected to be presented.

4.6.1.1 Data balancing method

The two cases described in Section 3.3.4 are evaluated, one where the training data is balanced by choosing random samples and one where the entire normal training data is clustered. By comparing the results in the general case, the cluster method results in a significant higher amount of false positives. When KDE is used, the cluster method generally results in a higher sensitivity with more true positives.

In the patient specific approach the result varies. For patient 36 and 39 the cluster method generally provides better performance, KNN with the cluster method for patient 39 actually gives a perfect result with all seizures found and zero false positives for several classifier setups. Similarly, patient 36 with cluster method will increase the performance when compared with the random approach. However, for both patient 21 and and 37 the clustering method results in decreased sensitivty, the amount of true positives is generally not within acceptable limits. Given the performances, the following results will be based on the random balancing method only. The complete result is found in Appendix.

4.6.1.2 Classification parameters

The initial performance evaluation indicate that ten features is marginally better than two or five features, however the differences are small and not always as suggested. A similar review indicates that KNN with two neighbors are significiantly worse than five or ten. Ten neighbors provide a marginally better specificity at the cost of a small reduction of sensitivity when compared with five neighbors. Random forest gives a better result with a larger amount of trees, however it is unclear whether 30 or 50 trees is better. SVM have a generally more stable result with linear kernel, the rbf kernel results varies too much to be considered reliable. The result does not depend on the evaluated box constraints.

Since the results are generally comparable and moderatly varying with change of method and parameters, it is not likely that a particulary good result have been found by accident. Roughly equal performance is achieved for all evaluated cases besides the kernel of SVM and KNN with two neighbors compared to five or more neighbors.

4.6.2 Post-processed example

The binary classifier output is processed as described in Section 3.5.2. The performance of the classifier will be presented in tables but the general effect of the post-processing method can be visualised using graphs. As an example, the highlighted result in the general column of Table 4.7 will be presented. It corresponds to KNN with ten neighbors, overlap integral reduction and veto 2 evaluated using patient 21. In Figure 4.16 the entire processed time of patient 21 is shown. The red lines corresponds to the estimated class where the normal class is represented by the value one and seizure is represented by two. The blue lines represents the actual seizures and the vector magnitude is shown to illustrate the activity. In this

resolution it is difficult to interpret details, however it is clear that the amount of red is significantly reduced, indicating a more stable result with less frequent change from seizure marking to non-seizure marking.

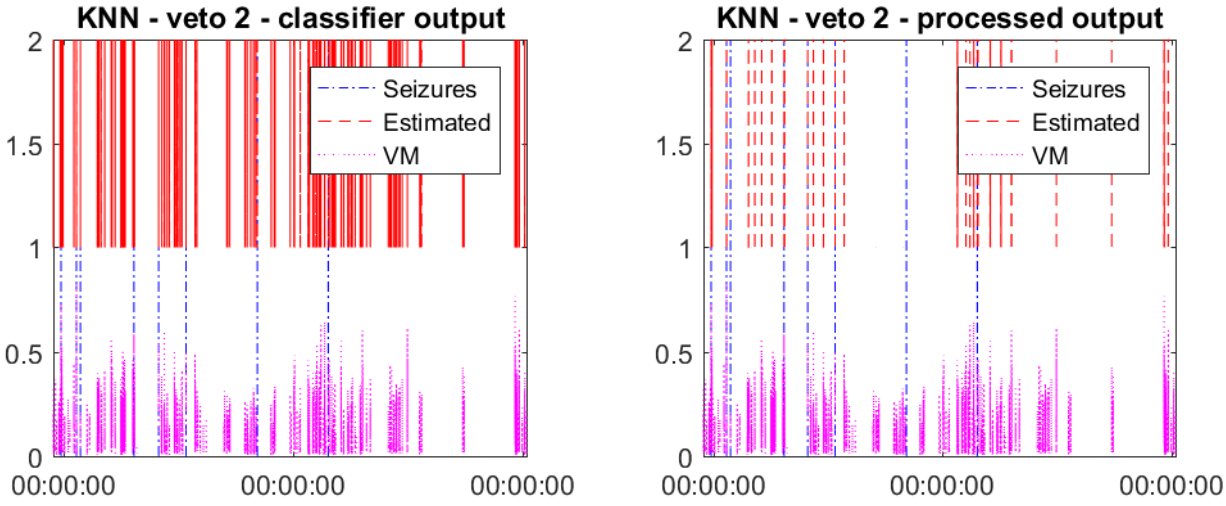

Figure 4.16: Left: Visualisation of the binary classification. Right: The corresponding post-processed classification.

The graphs in Figure 4.16 are shown using a shorter time window in Figure 4.17 and 4.18. The effect of the post-processing layer is evident when comparing the figures, over ten false positives have been reduced to two while the classified seizures remains. For this example, the complete post-processed performance without veto corresponds to 7 true positives and 111 false positives. By using veto 2 the amount of false positives is reduced to 31 at the cost of one true positive.

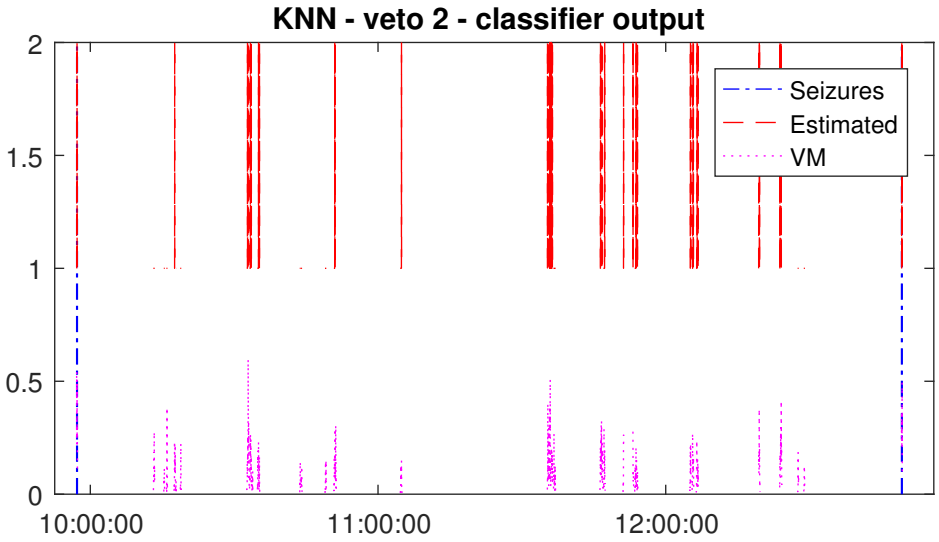

Figure 4.17: The binary classification example with frequent false positives.

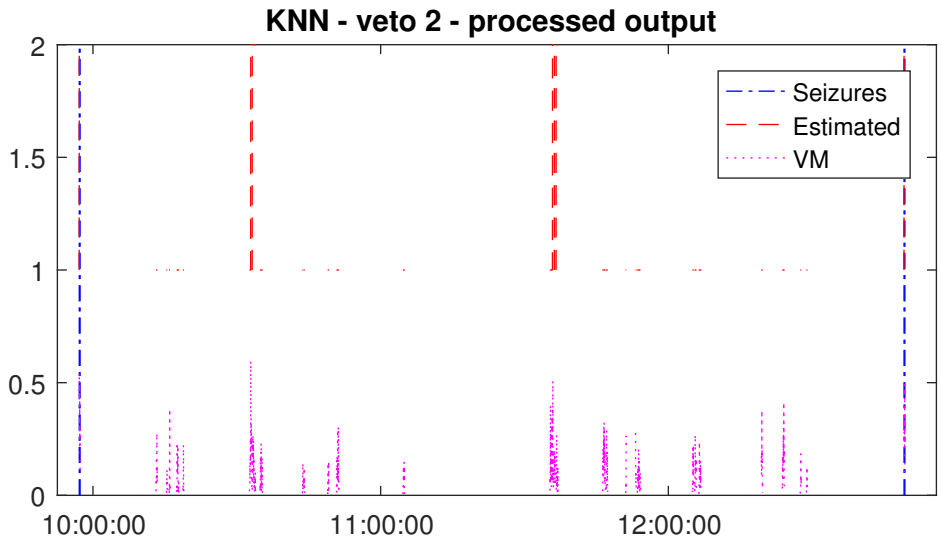

Figure 4.18: The post-processed example, signficicantly reduced amount of false positives.

4.6.3 Binary performance

The binary performance of the classification methods are evaluated with quarter based low activity data, 50-50 balanced training set and non-balanced test set. The binary performance is mainly evaluated to confirm that the classification method learns to recognize both seizure seconds and non-seizure seconds. Initially this performance was evaluated when training on the original imbalanced data set, however it was apparent that most classifiers only trained to recognize normal movements. The binary performance is found in Appendix B.

The results clearly indicate that all methods learn to recognize both classes. All combinations which includes KDE results in a significantly higher specificity while the sensitivity generally is reduced. A clear tradeoff is recognized when using KNN, if two neighbors are used the specificity is high while the sensitivity is small. When increasing neighbors these performances are shifted, the sensitivity is increased by reducing the specificity. The specificity of random forest generally increases with increased number of trees. The sensitivity varies, there is no clear indicator of whether the sensitivity is increased or decreased with the increased number of trees. SVM with rbf kernel generally have a much lower specificity than the linear kernel. In some cases the specificity is barely 50%, hence a large amount of false positives is to be expected. Nearly all combinations with the linear kernel have a specificity over 70%. Given the poor performance, the result of the rbf kernel is not shown.

4.6.4 Post-processed performance

As described in Section 2.6.2.2, the post-processed performance is presented using true positives (TP), false positives (FP) and false negatives (FN). The results tabu-

lated in this section will show the post-processed result for the original case, KDE, FS 10, OI 10 and one parameter for each classifier. Additionally, the original case without feature selection and KDE is evaluated to find seizures which often are missed, to verify the atypical and subtle seizures defined in Section 4.4. Hence, the amount of false negatives for each patient is summed over all feature dependent classifiers and their parameters and presented in one table for each patient. The amount of atypical false negatives encountered in the classification result for both the general and the patient specific approach is shown within parentheses accordingly.

Furthermore, to improve readability one result for each patient, approach and classification method will be highlighted using the color blue. The highlighted results are subjectively decided to be the "best" results aquired. For a result to be highlighed it must have the best specificity available given an acceptable sensitivity. The highlighted results may include false negatives if they are considered atypical, otherwise false negatives are not allowed. From the results which fulfill the false negative criterium, the one with the smallest amount of false positives is chosen, one for the general case and one for the patient specific case. Furthermore, the amount of atypical false negatives will be shown within parentheses.

4.6.4.1 Patient 21

Table 4.5 verifies the initial evaluation of the data set in Section 4.4, commonly missed seizures are seizure 3 and 7 which is not surprising as they are considered atypical.

Table 4.5: False negatives patient 21.

Seizure Nr	Date	No veto	Veto 1	Veto 2
1	03-Sep-2009 23:44:41	0	2	2
2	04-Sep-2009 01:21:18	2	9	8
3	04-Sep-2009 01:46:20	4	17	18
4	04-Sep-2009 07:22:02	2	2	5
5	04-Sep-2009 09:57:08	2	2	2
6	04-Sep-2009 12:49:09	2	2	2
7	04-Sep-2009 20:16:17	13	15	18
8	05-Sep-2009 03:41:41	3	3	8

The feature based classifiers of patient 21 have a higher sensitivity when the patient specific approach is used, the sensitivity is either higher or equal to the general case for all evaluated parameters. However, the false negatives in the general case are often atypical seizures and thus within acceptable limits. The patient specific approach generally leads to a significant increase of false positives, however for the best combinations (blue) the amount is arguably within similar range to the general case. Furthermore, the result indicates that veto 2 is superior in the patient specific approach while veto 1 is prefered in the general approach. The results also indicate

that OI10 is the best setup for patient 21. Furthermore, in Section 4.5, patient 21 was found not suitable for the KDE method which is confirmed by the performance.

The ANN performance is comparable with the feature based classifier performance. The patient specific performance is roughly equivalent while the general ANN approach generally have more false positives than the feature based classifiers.

Table 4.6: Post-processed results for random forest, patient 21.

Method: Random Forest, Trees = 30								
			General			Patient specific		
KDE	Red		No veto	Veto 1	Veto 2	No veto	Veto 1	Veto 2
		TP	6	6	5	7	7	7
		FP	57	36	20	101	70	53
		FN	2 (2)	2 (2)	3 (2)	1 (1)	1 (1)	1 (1)
✗		TP	2	1	1	2	2	2
		FP	44	11	3	44	27	20
		FN	6 (2)	7 (2)	7 (2)	6 (2)	6 (2)	6 (2)
	FS 10	TP	7	6	5	8	7	7
		FP	86	58	37	101	73	53
		FN	1 (1)	2 (2)	3 (2)	0	1 (1)	1 (1)
	OI 10	TP	7	5	5	8	8	8
		FP	65	38	19	102	73	53
		FN	1 (1)	3 (2)	3 (2)	0	0	0

Table 4.7: Post-processed results for KNN, patient 21.

Method: KNN, Neighbors = 10								
			General			Patient specific		
KDE	Red		No veto	Veto 1	Veto 2	No veto	Veto 1	Veto 2
		TP	7	6	6	7	7	7
		FP	115	53	36	81	63	48
		FN	1 (1)	2 (2)	2 (2)	1 (1)	1 (1)	1 (1)
✗		TP	2	1	1	3	3	3
		FP	66	17	5	37	25	13
		FN	6 (2)	7 (2)	7 (2)	5 (2)	5 (2)	5 (2)
	FS 10	TP	7	6	5	8	8	6
		FP	104	59	38	83	62	44
		FN	1 (1)	2 (2)	3 (2)	0	0	2 (2)
	OI 10	TP	7	6	6	7	7	7
		FP	111	53	31	83	64	54
		FN	1 (1)	2 (2)	2 (2)	1 (1)	1 (1)	1 (1)

Table 4.8: Post-processed results for logistic regression, patient 21.

Method: Logistic Regression								
			General			Patient specific		
KDE	Red		No veto	Veto 1	Veto 2	No veto	Veto 1	Veto 2
		TP	7	5	4	8	8	7
		FP	76	49	25	167	85	58
		FN	1 (1)	3 (2)	4 (2)	0	0	1 (0)
✗		TP	4	2	1	4	3	3
		FP	60	11	4	66	42	31
		FN	4 (2)	6 (2)	7 (2)	4 (2)	5 (2)	5 (2)
	FS 10	TP	6	4	4	7	6	6
		FP	79	44	26	150	85	59
		FN	2 (1)	4 (2)	4 (2)	1 (1)	2 (2)	2 (2)
	OI 10	TP	7	6	4	7	7	7
		FP	82	40	26	117	82	58
		FN	1 (1)	2 (2)	4 (2)	1 (1)	1 (1)	1 (1)

Table 4.9: Post-processed results for SVM, patient 21.

Method: SVM, Kernel = linear, box constraint = 1								
			General			Patient specific		
KDE	Red		No veto	Veto 1	Veto 2	No veto	Veto 1	Veto 2
		TP	7	5	4	8	7	6
		FP	82	38	21	117	81	56
		FN	1 (1)	3 (2)	4 (2)	0	1 (1)	2 (2)
✗		TP	3	1	0	3	3	2
		FP	58	9	2	46	29	19
		FN	5 (2)	7 (2)	8 (2)	5 (2)	5 (2)	6 (2)
	FS 10	TP	7	5	5	7	7	6
		FP	77	39	25	110	79	58
		FN	1 (1)	3 (2)	3 (2)	1 (0)	1 (0)	2 (1)
	OI 10	TP	7	6	4	7	7	7
		FP	92	42	25	97	70	55
		FN	1 (1)	2 (2)	4 (2)	1 (1)	1 (1)	1 (1)

Table 4.10: Post-processed results for ANN, patient 21.

Method: ANN								
			General			Patient specific		
H neurons	H depth		No veto	Veto 1	Veto 2	No veto	Veto 1	Veto 2
0	0	TP	7	6	6	5	5	5
		FP	146	73	59	118	69	58
		FN	1 (1)	2 (2)	2 (2)	3 (2)	3 (2)	3 (2)
2	1	TP	7	6	6	5	5	5
		FP	180	86	72	140	96	58
		FN	1 (1)	2 (1)	2 (2)	3 (2)	3 (2)	3 (2)
2	2	TP	6	6	4	6	6	6
		FP	107	56	42	145	78	68
		FN	2 (2)	2 (2)	4 (2)	2 (2)	2 (2)	2 (2)
5	1	TP	6	6	3	6	6	5
		FP	110	66	48	97	56	48
		FN	2 (2)	2 (2)	5 (2)	2 (2)	2 (2)	3 (2)
5	2	TP	6	6	4	6	6	6
		FP	120	62	52	123	69	56
		FN	2 (2)	2 (2)	4 (2)	2 (2)	2 (2)	2 (2)
10	1	TP	6	6	6	6	6	5
		FP	148	71	58	104	65	53
		FN	2 (2)	2 (2)	2 (2)	2 (2)	2 (2)	3 (2)
10	2	TP	6	6	4	7	7	5
		FP	129	73	59	83	59	53
		FN	2 (2)	2 (2)	4 (2)	1 (1)	1 (1)	3 (2)
18	1	TP	6	6	4	6	6	5
		FP	111	60	46	97	61	56
		FN	2 (2)	2 (2)	4 (2)	2 (2)	2 (2)	3 (2)
18	2	TP	6	6	5	7	7	6
		FP	113	66	52	94	64	56
		FN	2 (2)	2 (2)	3 (2)	1 (1)	1 (1)	2 (2)

4.6.4.2 Patient 36

Patient 36 have no atypical seizures according to Section 4.4, which is verified by Table 4.11. This implies that all seizures should be accurately classified for it to be acceptable. Additionally, the amount of seizures is very small, the patient specific approach only have two seizures in the training set.

Table 4.11: False negatives patient 36.

Seizure Nr	Date	No veto	Veto 1	Veto 2
1	05-Nov-2010 06:06:59	2	2	2
2	07-Nov-2010 01:54:28	2	2	2
3	07-Nov-2010 06:11:04	2	2	2

Classification of patient 36 have a high true positive rate, regardless of whether patient specific or general approach is used. The amount of false positives is generally reduced when the patient specific approach is used, indicating that it might be preferable. In the patient specific case, KDE combined with no veto is prefered for all methods but random forest. This is not surprising given the results of Section 4.5 which clearly indicated that the patient specific KDE was efficient for patient 36. The general approach is improved by using veto 2 which reduces the amount of false positivies significantly.

The ANN performance of the general approach results in a smaller amount of false positives than the feature based classifiers while the opposite relation is aquired using the patient specific approach.

Table 4.12: Post-processed results for random forest, patient 36.

Method: Random Forest, Trees = 30								
			General			Patient specific		
KDE	Red		No veto	Veto 1	Veto 2	No veto	Veto 1	Veto 2
		TP	3	3	3	3	3	3
		FP	114	72	58	93	71	52
		FN	0	0	0	0	0	0
✗		TP	2	2	2	1	1	1
		FP	95	45	30	17	9	4
		FN	1	1	1	2	2	2
	FS 10	TP	3	3	3	3	3	3
		FP	118	80	64	107	73	47
		FN	0	0	0	0	0	0
	OI 10	TP	3	3	3	3	3	3
		FP	124	80	63	93	69	53
		FN	0	0	0	0	0	0

Table 4.13: Post-processed results for KNN, patient 36.

Method: KNN, Neighbors = 10								
			General			Patient specific		
KDE	Red		No veto	Veto 1	Veto 2	No veto	Veto 1	Veto 2
		TP	3	3	3	3	3	3
		FP	120	81	62	100	61	36
		FN	0	0	0	0	0	0
✗		TP	2	2	1	3	2	2
		FP	93	47	28	24	13	8
		FN	1	1	2	0	1	1
	FS 10	TP	3	3	3	3	3	3
		FP	112	78	59	132	65	40
		FN	0	0	0	0	0	0
	OI 10	TP	3	3	3	3	3	3
		FP	126	82	57	97	65	44
		FN	0	0	0	0	0	0

Table 4.14: Post-processed results for logistic regression, patient 36.

Method: Logistic Regression								
			General			Patient specific		
KDE	Red		No veto	Veto 1	Veto 2	No veto	Veto 1	Veto 2
		TP	3	3	3	3	3	3
		FP	117	90	71	127	75	44
		FN	0	0	0	0	0	0
✗		TP	2	2	2	3	2	2
		FP	95	40	27	25	16	9
		FN	1	1	1	0	1	1
	FS 10	TP	3	3	3	3	3	3
		FP	119	87	64	126	73	37
		FN	0	0	0	0	0	0
	OI 10	TP	3	3	3	3	3	3
		FP	119	89	73	116	76	53
		FN	0	0	0	0	0	0

Table 4.15: Post-processed results for SVM, patient 36.

Method: SVM, Kernel = linear, box constraint = 1								
			General			Patient specific		
KDE	Red		No veto	Veto 1	Veto 2	No veto	Veto 1	Veto 2
		TP	3	3	3	3	3	3
		FP	111	86	76	94	61	36
		FN	0	0	0	0	0	0
✗		TP	1	1	1	3	2	2
		FP	80	34	22	20	14	10
		FN	2	2	2	0	1	1
	FS 10	TP	3	3	3	3	3	3
		FP	126	90	79	87	55	35
		FN	0	0	0	0	0	0
	OI 10	TP	3	3	3	3	3	3
		FP	135	92	77	86	64	46
		FN	0	0	0	0	0	0

Table 4.16: Post-processed results for ANN, patient 36.

Method: ANN								
			General			Patient specific		
H neurons	H depth		No veto	Veto 1	Veto 2	No veto	Veto 1	Veto 2
0	0	TP	3	3	3	3	3	3
		FP	69	69	57	138	91	72
		FN	0	0	0	0	0	0
2	1	TP	3	3	3	3	3	3
		FP	111	92	79	125	98	90
		FN	0	0	0	0	0	0
2	2	TP	2	2	2	2	2	2
		FP	56	54	44	76	64	58
		FN	1 (0)	1 (0)	1 (0)	1 (0)	1 (0)	1 (0)
5	1	TP	3	3	3	3	3	2
		FP	71	66	57	110	85	69
		FN	0	0	0	0	0	1 (0)
5	2	TP	3	3	3	3	3	3
		FP	92	82	75	114	86	76
		FN	0	0	0	0	0	0
10	1	TP	3	3	3	2	2	2
		FP	70	67	58	79	57	48
		FN	0	0	0	1 (0)	1 (0)	1 (0)
10	2	TP	3	3	3	2	2	2
		FP	60	57	42	84	64	55
		FN	0	0	0	1 (0)	1 (0)	1 (0)
18	1	TP	3	3	3	2	2	2
		FP	81	75	66	105	77	68
		FN	0	0	0	1 (0)	1 (0)	1 (0)
18	2	TP	3	3	3	3	3	3
		FP	56	55	49	96	77	63
		FN	0	0	0	0	0	0

4.6.4.3 Patient 37

According to Section 4.4, the atypical seizures are 6, 15, 16, 17, 18 and 20. Table 4.17 show that these are commonly missed in the classification process, verifying that they are too dissimilar or subtle to detect. In the following tables, the amount of false negatives which are atypical is shown within parentheses. Additionally, given the result seizure 19 is quite commonly missed as well, however as it was not initially defined as atypical it will not be shown within parentheses.

Both in the patient specific approach and the general approach, the best results are most commonly aquired using the setup FS 10. In the patient specific case, the setup

is usually combined with veto 1 for best result while the general approach benefits from veto 2. Furthermore, KDE is not preferable when evaluating patient 37. When using KDE the amount of false negatives is severly increased, and corresponds to significant amount of seizures which are not considered as atypical. This supports the results of Section 4.5, KDE was not expected to be efficient for patient 37.

Generally the patient specific approach results in a smaller amount of true positives than the general approach. However this reduction is often within the false negative criterium, the additional false negatives are often atypical. The amount of false positives are similar in both approaches. The patient specific ANN results in more false negatives than the general approach. The highlighted patient specific result does not fulfill the false negative criterium, it contains two false negatives which were not specified as atypical seizure. However, the atypical seizures were confirmed as such by evaluating the response of the feature based classifiers, which indicates that it might not be directly transferable to the ANN approach which mainly depends on accelerometer data.

Table 4.17: False negatives patient 37.

Seizure Nr	Date	No veto	Veto 1	Veto 2
1	09-Feb-2011 22:55:38	2	2	2
2	09-Feb-2011 23:40:06	2	2	2
3	10-Feb-2011 01:02:46	2	2	2
4	10-Feb-2011 01:06:45	2	2	2
5	10-Feb-2011 01:14:52	1	1	2
6	10-Feb-2011 01:37:02	0	9	20
7	10-Feb-2011 01:46:07	1	2	2
8	10-Feb-2011 03:50:42	0	4	7
9	10-Feb-2011 04:55:43	1	2	2
10	10-Feb-2011 06:04:59	2	2	2
11	10-Feb-2011 08:48:24	2	2	2
12	10-Feb-2011 14:51:15	3	3	8
13	10-Feb-2011 22:30:39	2	2	2
14	10-Feb-2011 23:53:50	2	5	5
15	11-Feb-2011 00:21:56	2	13	22
16	11-Feb-2011 00:29:06	11	22	22
17	11-Feb-2011 00:34:05	4	17	22
18	11-Feb-2011 00:41:02	6	20	22
19	11-Feb-2011 02:01:03	4	6	7
20	11-Feb-2011 04:10:09	2	11	17
21	11-Feb-2011 04:21:17	1	2	4
22	11-Feb-2011 04:50:03	2	2	2
23	11-Feb-2011 06:09:15	1	2	2
24	11-Feb-2011 07:33:48	2	2	2

Table 4.18: Post-processed results for random forest, patient 37.

Method: Random Forest, Trees = 30								
			General			Patient specific		
KDE	Red		No veto	Veto 1	Veto 2	No veto	Veto 1	Veto 2
		TP	22	21	18	22	18	18
		FP	82	55	37	79	45	27
		FN	2 (2)	3 (3)	6 (6)	2 (1)	6 (5)	6 (5)
✗		TP	17	8	6	16	11	9
		FP	67	20	16	57	23	15
		FN	7 (0)	16 (6)	18 (6)	8 (3)	13 (5)	15 (6)
	FS 10	TP	23	21	17	24	17	15
		FP	100	61	44	96	56	37
		FN	1 (1)	3 (3)	7 (6)	0	7 (6)	9 (6)
	OI 10	TP	22	21	18	23	18	18
		FP	80	61	39	76	47	29
		FN	2 (2)	3 (3)	6 (6)	1	6 (5)	6 (5)

Table 4.19: Post-processed results for KNN, patient 37.

Method: KNN, Neighbors = 10								
			General			Patient specific		
KDE	Red		No veto	Veto 1	Veto 2	No veto	Veto 1	Veto 2
		TP	24	20	17	24	20	16
		FP	84	65	37	90	53	30
		FN	0	4 (4)	7 (6)	0	4 (4)	8 (6)
✗		TP	17	8	4	15	12	7
		FP	52	15	12	66	27	13
		FN	7 (0)	16 (4)	20 (6)	9 (3)	12 (5)	17 (6)
	FS 10	TP	23	21	18	23	19	17
		FP	89	56	35	90	52	28
		FN	1 (1)	3 (3)	6 (6)	1 (1)	5 (5)	7 (6)
	OI 10	TP	24	22	17	23	21	17
		FP	78	59	31	92	50	25
		FN	0	2 (2)	7 (6)	1 (1)	3 (3)	7 (6)

Table 4.20: Post-processed results for logistic regression, patient 37.

Method: Logistic Regression								
			General			Patient specific		
KDE	Red		No veto	Veto 1	Veto 2	No veto	Veto 1	Veto 2
		TP	22	22	20	24	18	16
		FP	93	69	47	93	53	32
		FN	2 (2)	2 (2)	4 (4)	0	6 (5)	8 (6)
✗		TP	16	6	7	12	8	7
		FP	67	25	16	68	26	18
		FN	8 (0)	18 (6)	17 (6)	12 (3)	16 (5)	17 (6)
	FS 10	TP	22	21	18	24	19	16
		FP	84	59	45	96	49	34
		FN	2 (2)	3 (3)	6 (6)	0	5 (5)	8 (6)
	OI 10	TP	21	21	19	22	17	16
		FP	81	64	51	100	51	36
		FN	3 (3)	3 (3)	5 (5)	2 (1)	7 (5)	8 (6)

Table 4.21: Post-processed results for SVM, patient 37.

Method: SVM, Kernel = linear, box constraint = 1								
			General			Patient specific		
KDE	Red		No veto	Veto 1	Veto 2	No veto	Veto 1	Veto 2
		TP	21	20	18	24	17	15
		FP	65	52	37	100	53	35
		FN	3 (3)	4 (3)	6 (5)	0	7 (5)	9 (6)
✗		TP	13	4	4	13	10	7
		FP	56	25	12	72	23	17
		FN	11 (1)	20 (5)	20 (5)	11 (2)	14 (4)	17 (6)
	FS 10	TP	22	20	15	23	20	16
		FP	89	62	42	102	54	32
		FN	2 (2)	4 (4)	9 (6)	1 (1)	4 (4)	8 (6)
	OI 10	TP	21	19	17	23	19	15
		FP	69	54	43	108	54	36
		FN	3 (3)	5 (4)	7 (5)	1 (1)	5 (3)	9 (6)

Table 4.22: Post-processed results for ANN, patient 37.

Method: ANN								
			General			Patient specific		
H neurons	H depth		No veto	Veto 1	Veto 2	No veto	Veto 1	Veto 2
0	0	TP	21	21	21	18	18	14
		FP	115	91	72	74	54	44
		FN	3 (3)	3 (3)	3 (3)	6 (4)	6 (4)	10 (5)
2	1	TP	21	21	20	18	18	15
		FP	96	79	59	85	60	49
		FN	3 (3)	3 (3)	4 (3)	6 (4)	6 (4)	9 (4)
2	2	TP	22	22	22	17	17	16
		FP	115	92	71	95	66	54
		FN	2 (2)	2 (2)	2 (2)	7 (3)	7 (3)	8 (4)
5	1	TP	21	21	21	18	16	16
		FP	90	80	59	77	65	49
		FN	3 (3)	3 (3)	3 (3)	6 (3)	8 (4)	8 (4)
5	2	TP	21	21	20	18	17	17
		FP	90	74	64	87	57	43
		FN	3 (3)	3 (3)	4 (4)	6 (3)	7 (4)	7 (4)
10	1	TP	22	22	22	17	17	16
		FP	105	91	69	66	56	44
		FN	2 (2)	2 (2)	2 (2)	7 (4)	7 (4)	8 (5)
10	2	TP	21	21	21	18	18	15
		FP	100	86	68	62	53	38
		FN	3 (2)	3 (2)	3 (2)	6 (4)	6 (4)*	9 (5)
18	1	TP	22	22	22	16	16	14
		FP	103	99	79	67	57	47
		FN	2 (2)	2 (2)	2 (2)	8 (4)	8 (4)	10 (5)
18	2	TP	22	22	22	17	17	15
		FP	124	109	87	58	53	40
		FN	2 (2)	2 (2)	2 (2)	7 (4)	7 (4)	9 (5)
* Does not fulfill false negative criterium								

4.6.4.4 Patient 39

Patient 39 have no atypical seizure which is verified in Table 4.23. Additionally, the amount of seizures is very small, the patient specific approach only have one seizure in the training set. The patient specific approach for both the feature based methods and ANN is evaluated nonetheless, however it is important to note the limitations of the data set. The ANN models will be trained using only 75% of one seizure, which is not much to go on.

Table 4.23: False negatives patient 39.

Seizure Nr	Date	No veto	Veto 1	Veto 2
1	24-Mar-2011 17:36:53	3	3	3
2	24-Mar-2011 21:30:02	2	2	2

The results in Section 4.5 clearly indicated the possible efficiency of KDE for patient 39 which is confirmed in the final performance, in the general approach the best results are aquired using KDE and veto 2. However, the general approach results in a large amount of false positives, which is significantly reduced using the patient specific approach. Additionally, the patient specific approach with KNN results in the best performance of all evaluated setups which indicate clearly that the patient specific approach is optimal for patient 39. ANN results in high sensitivity for both the general approach and patient specific approach. The amount of false positives is significantly smaller when the patient specific approach is used. However as previously mentioned, it is important to note that the patient specific approach only have one seizure in the training fold.

Table 4.24: Post-processed results for random forest, patient 39.

Method: Random Forest, Trees = 30								
			General			Patient specific		
KDE	Red		No veto	Veto 1	Veto 2	No veto	Veto 1	Veto 2
		TP	2	2	2	2	2	2
		FP	80	73	64	41	36	34
		FN	0	0	0	0	0	0
✗		TP	2	2	2	1	1	1
		FP	78	40	44	16	7	4
		FN	0	0	0	1	1	1
	FS 10	TP	2	2	2	2	2	2
		FP	77	66	59	55	48	40
		FN	0	0	0	0	0	0
	OI 10	TP	2	2	2	2	2	2
		FP	82	75	61	47	39	38
		FN	0	0	0	0	0	0

Table 4.25: Post-processed results for KNN, patient 39.

Method: KNN, Neighbors = 10								
			General			Patient specific		
KDE	Red		No veto	Veto 1	Veto 2	No veto	Veto 1	Veto 2
		TP	2	2	2	2	2	2
		FP	85	75	57	6	4	4
		FN	0	0	0	0	0	0
✗		TP	1	1	1	0	0	0
		FP	77	47	42	1	0	0
		FN	1	1	1	2	2	2
	FS 10	TP	2	2	2	2	2	2
		FP	81	72	59	6	4	5
		FN	0	0	0	0	0	0
	OI 10	TP	2	2	2	2	2	2
		FP	83	69	60	13	9	7
		FN	0	0	0	0	0	0

Table 4.26: Post-processed results for logistic regression, patient 39.

Method: Logistic Regression								
			General			Patient specific		
KDE	Red		No veto	Veto 1	Veto 2	No veto	Veto 1	Veto 2
		TP	2	2	2	1	1	1
		FP	78	70	62	82	58	22
		FN	0	0	0	1	1	1
✗		TP	2	2	2	1	1	1
		FP	79	46	43	21	6	5
		FN	0	0	0	1	1	1
	FS 10	TP	2	2	2	2	2	1
		FP	79	69	62	89	54	28
		FN	0	0	0	0	0	1
	OI 10	TP	2	2	2	2	2	2
		FP	86	70	61	51	42	29
		FN	0	0	0	0	0	0

Table 4.27: Post-processed results for SVM, patient 39.

Method: SVM, Kernel = linear, box constraint = 1								
			General			Patient specific		
KDE	Red		No veto	Veto 1	Veto 2	No veto	Veto 1	Veto 2
		TP	2	2	2	2	2	2
		FP	78	71	62	32	26	20
		FN	0	0	0	0	0	0
✗		TP	2	2	2	0	0	0
		FP	79	47	45	9	5	4
		FN	0	0	0	2	2	2
	FS 10	TP	2	2	2	2	2	2
		FP	82	69	66	46	40	31
		FN	0	0	0	0	0	0
	OI 10	TP	2	2	2	2	2	2
		FP	85	67	58	32	28	23
		FN	0	0	0	0	0	0

Table 4.28: Post-processed results for ANN, patient 39.

Method: ANN								
			General			Patient specific		
H neurons	H depth		No veto	Veto 1	Veto 2	No veto	Veto 1	Veto 2
0	0	TP	2	2	2	1	1	1
		FP	72	68	59	27	24	19
		FN	0	0	0	1 (0)	1 (0)	1 (0)
2	1	TP	2	2	2	2	2	2
		FP	74	71	64	27	27	17
		FN	0	0	0	0	0	0
2	2	TP	2	2	2	2	2	2
		FP	77	71	60	30	27	21
		FN	0	0	0	0	0	0
5	1	TP	2	2	2	2	2	2
		FP	70	68	63	32	32	25
		FN	0	0	0	0	0	0
5	2	TP	2	2	2	2	2	2
		FP	84	79	72	41	41	33
		FN	0	0	0	0	0	0
10	1	TP	2	2	2	1	1	1
		FP	75	69	62	21	21	15
		FN	0	0	0	1 (0)	1 (0)	1 (0)
10	2	TP	2	2	2	2	2	2
		FP	83	79	72	49	48	38
		FN	0	0	0	0	0	0
18	1	TP	2	2	2	2	2	2
		FP	66	65	59	30	30	18
		FN	0	0	0	0	0	0
18	2	TP	2	2	2	2	2	2
		FP	77	75	66	40	39	30
		FN	0	0	0	0	0	0

4.6.5 Summary

All highlighted performances are summarized in Table 4.29, where the false positives is displayed as the amount of false positives per hour. The false negative criterium is fulfilled for all combinations except the ANN, patient specific approach for patient 37. However the amount of false negatives are less than the total number of atypical seizures.

Since the amount of true positives is sufficient for each combination, the preferable approach can be chosen such that the amount of false alarms are at a minimum. The results indicate that the general approach is superior for patient 21 and 37 while

the patient specific approach is preferable for patient 36 and 39.

Table 4.29: Summary of post-processed results.

Patient			randf	KNN	logr	SVM	ANN
21	Gen	TP	6	6	6	6	6
		FP/h	1.36	1.17	1.51	2.11	2.11
		FN	2	2	2	2	2
21	Spec	TP	8	6	7	7	6
		FP/h	2.00	1.66	2.19	2.08	2.11
		FN	0	2	1	1	2
36	Gen	TP	3	3	3	3	3
		FP/h	1.74	1.71	1.92	2.32	1.26
		FN	0	0	0	0	0
36	Spec	TP	3	3	3	3	3
		FP/h	1.41	0.72	0.75	0.60	1.90
		FN	0	0	0	0	0
37	Gen	TP	18	18	18	20	21
		FP/h	0.74	0.70	0.90	1.24	1.18
		FN	6	6	6	4	3
37	Spec	TP	24	21	19	20	18
		FP/h	1.92	1.00	0.98	1.08	1.06
		FN	0	3	5	4	6*
39	Gen	TP	2	2	2	2	2
		FP/h	2.29	3.26	2.46	2.58	3.37
		FN	0	0	0	0	0
39	Spec	TP	2	2	2	2	2
		FP/h	1.94	0.23	1.66	1.14	0.97
		FN	0	0	0	0	0
* Does not fulfill false negative criterium							

To evaluate which setup results in the best performance, Table 4.29 can be modified by rating each performance. Since the true positives are within acceptable limits for each setup except one, the rating will be based on the amount of false positives. Each item in each row is graded between one and five, where one corresponds to the best achieved performance for the current patient and approach, and five corresponds to the worst performance. Each column, corresponding to classification methods, is summarized to give a total rating for each method, and a low total rating indicates a classifier method with good performance. The ratings are shown in Table 4.30. The ratings indicates that KNN is the superior classification method while the other methods results in similar ratings.

Table 4.30: Rating of the summary of Table 4.29.

Patient		randf	KNN	logr	SVM	ANN
21	Gen	2	1	3	4	4
21	Spec	2	1	5	3	4
36	Gen	3	2	4	5	1
36	Spec	4	2	3	1	5
37	Gen	2	1	3	5	4
37	Spec	5	2	1	4	3
39	Gen	1	4	2	3	5
39	Spec	5	1	4	3	2
	Total:	24	14	25	28	28

4.7 Multi-class Features

As mentioned in 1.3.2.1, patient 21 have both considered seizure types. Hence, the patient is evaluated to indicate differences between HMS, GTCS and non-seizure movements. Figures 4.19 and 4.20 show that GTCS has got a significantly smaller overlap with normal data than HMS, indicating that it might be easier to accurately identify GTCS than HMS. The figures also show a significant difference between GTCS and HMS, indicating that the current set of features could be used for multi-class classification. The newly introduced feature Spectral Edge Frequency 95 is especially interesting for multi-class classification as it seems to provide a significant difference in distribution. Based on the histograms, it is definately possible to evaluate this feature set as a basis for multi-class classification in future work.

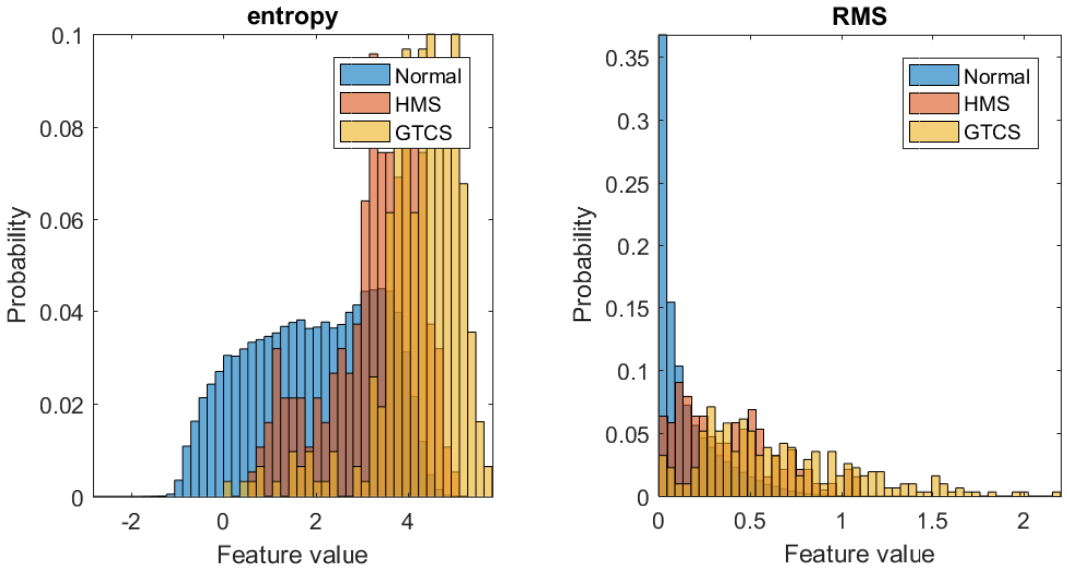

Figure 4.19: Left: Entropy of patient 21. Right: RMS of patient 21.

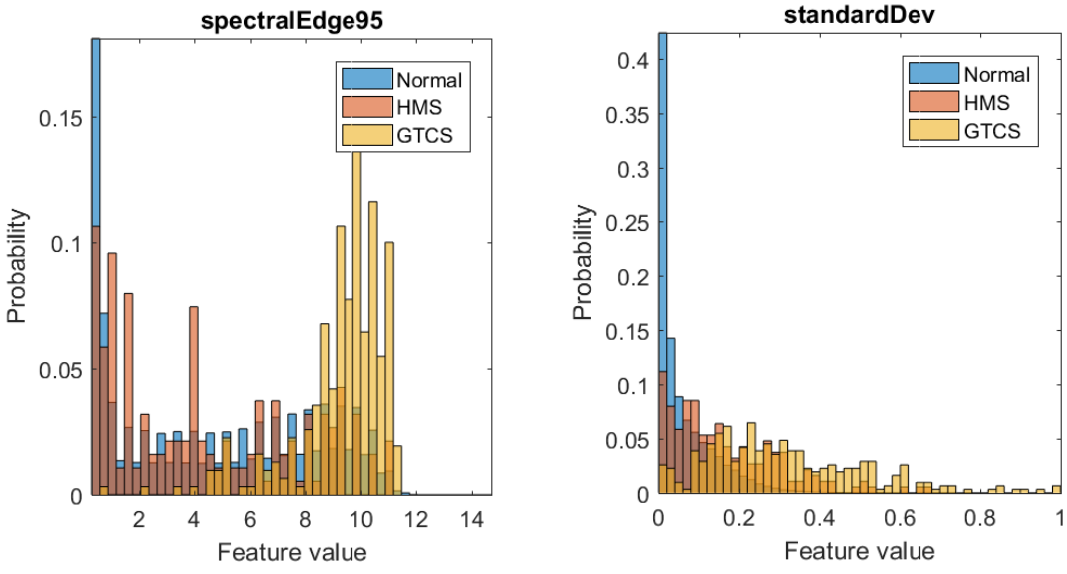

Figure 4.20: Left: Spectral Edge Frequency 95 of patient 21. Right: Standard deviation of patient 21.

5
Discussion

5.1 Data set

The initial evaluation of the accelerometer data was briefly discussed in Section 1.3.2. There is a wide range of seizure manifestations where some are very subtle or atypical while others are not. As discussed, this was expected to have a negative impact on the classifier performance. This was later confirmed by the performance statistics in Section 4.6.4. The atypical seizures were often misclassified, indicating that the seizures were too dissimilar to the rest of the available data.

There are several possible reasons as to why some seizures were not accurately classified. One possibility is that these seizures are not representative of HMS. The original seizure onsets were determined by video-EEG, the physicians did not validate the accelerometer data for each seizure. Perhaps the motoric response of these seizures are not significant enough for them to actually be considered as typical HMS. One other possibility is that the seizures are valid HMS, but due to the limited data set the variability is not sufficiently represented. Hence, some seizures which are valid, but too different from the rest are not recognized as HMS.

For future work, each seizure and its corresponding accelerometer data should be evaluated by physicians to determine whether the motor manifestations are significant enough to represent HMS. While all current seizures are valid HMS according to EEG, a stricter selection where the severity of motor manifestations are taken into account would make the data more suitable for the accelerometry based machine learning approach. Discussion with physicians and the results in Section 4.4 initiates this work, a set of subtle and atypical seizures have been identified. In the result presented in this thesis these seizures are included, but an initial step in future work could be to evaluate the performance when these seizures have been removed. Hopefully this will improve the performance by reducing the amount of false positives. Additionally, more data should be collected to increase the variability covered by the classifiers.

5.2 Data reduction

As mentioned in Section 3.3 an imbalanced dataset can lead to poorly trained classifiers with the main focus to accurately classify only the dominant class. The different methods used for data reduction have been used with varying result.

5.2.1 Low-activity build

The activity based reduction based on either quarter activity or hour activity provides an important initial data reduction. Upon review the hour based reduction is not sufficient to describe the low activity of a patient, only patient 37 have low activity hours with seizures which can be seen in Table 4.1. Additionally, when considering only complete hours there may be low-activity movements which are overlooked and high-activity movements which are kept. For instance if there is an low-activity hour which ends with the patient waking up the entire hour might be discarded. To that end, only the data reduction based on quarter activity is used. In Table 4.2 it is evident that there are low-activity seizures for all patients but patient 39. The low-activity build is constructed to simulate nocturnal activity. However, the resulting build is much less restricted and controlled when compared to the literature. In the literature only actual sleep is used, in the build of this thesis general low activity is considered which might be during the day as well.

All seizure data is kept regardless of activity. One can argue that the behavior before, during and after might be different depending on whether the patient is asleep or not, however given the small amount of seizure data the potential differences have to be accepted.

5.2.2 Balancing dataset

If the training dataset is balanced by random sampling there is no guarantee that the data covers the range of all types of normal movements. However, under the assumption that the normal movements follow some kind of distribution it is likely that a set of random samples will sufficiently describe the distribution. If the clustering method is used instead, the normal data will be based on all available normal data. It is however important to note that the clustering method will result in classifiers which have not been trained on original data. The clustering method clusters data points in close proximity to each other, and the centroid of the cluster is the new data point. This averaging approach will generally reduce the amount of normal outliers in the training data since the means will be placed closer to the dominating clusters. As a consequence, the classifiers will not be trained on the extreme values presented in the original normal data set. Naturally this can result in an increased amount of false positives, since normal outliers of the test set might not have been represented in the training set.

The reasoning above is supported by the results. As briefly discussed in Section 4.6.1.1, the general performance with the clustering method results in a larger amount of false positives. In the patient specific approach, the performance of the clustering method varies among the patients. For patients with limited separability the amount of true positives is decreased and generally not within acceptable limits. For the remaining patients where the seizures are further away from the dominating normal clusters the performance is significantly improved, resulting in the best performance encountered.

5.3 Features

The histograms in Section 4.3 shows that there are differences in the distributions of normal movements and seizure movements which indicates that classification is possible. However, the significant overlap between the two classes implies that it will be difficult to achieve both high sensitivity and high specificity using the available data set. This is supported by the cross-validation and post-processed statistics, presented in Section 4.6.4; the classifiers are clearly learning to identify seizures but at a cost of specificity. Additionally, the current feature set differentiates HMS and GTCS, hence it can be used as a basis for multi-class classification.

Additionally, since the performance of the feature based classifiers is better than that of the ANN, it appears that the current feature set successfully represents the characteristics of the data set.

5.4 Performance

5.4.1 KDE

The KDE reduction of the data sets significantly increases the specificity in all evaluated cases. However this is usually at a cost of sensitivity. The KDE reduction will remove seizures if the test patients seizure movements are very similar to the normal movements of the patients used in the training process, hence subtle seizures risk being removed. This risk might be reduced if a stricter seizure selection is considered, as described in Section 5.1.

As mentioned in Section 4.5, KDE is accurate enough to be used as the sole classification algorithm for patient 39 and nearly as efficient for patient 36. For patient 21 and 37 the method is not sufficient, the feature overlap for seizures and non-seizures are too significant. Hence the KDE method described in literature is not guaranteed to work, some kind of improvement must be done in order for it to function efficiently regardless of patient. The performance can be improved by removing the specified atypical and subtle seizures, but it will not be perfect regardless.

Additionally, the KDE algorithm could be extended to a multivariate method which covers more dimensions than the currently used bivariate case. However, the brute-force approach evaluated in this thesis would result in a significantly increased computation time if the number of dimensions are increased.

5.4.2 Post-processed statistics

The veto is chosen to depend on correlation and entropy after inspection of the histograms in Section 4.3. There is a clear difference in the distributions of normal movements and seizure movements. The additional feature frequency peak is only used to disable the veto; when there is an unusually high frequency peak it is most

likely not a normal movement. All features and feature combinations have not been tested, it is possible that the result can be improved further.

Regardless of method there is a high amount of false positives, however 1 FP/h is not uncommon for continuous monitoring. This can partly be explained by all the previous discussions, the less restricted low-activity build, the atypical and subtle seizures and the significant overlap in the feature space. It is also important to note that the training of the models is performed using a balanced dataset while the test set is not. By balancing the training set neither normal movements nor seizure movements is prefered during the training process. But the test set does not follow the balanced distribution, there is more normal data than seizure data. As a consequence even though the binary classification specficity is roughly equivalent to the sensitivity there will be more misclassified normal seconds than seizure seconds. When the same testing was performed with non-balanced training set, the classifiers basically ignored the seizures and had all focus on the specificity.

Given the rating found in Table 4.30, it appears that KNN is the superior method. KNN also gave good results for detection of GTCS [9], which indicates that the method is promising for multi-class classification as well.

5.5 General vs Patient specific approach

The performance does not fully indicate which approach is preferable. As previously mentioned, the general approach can be considered superior for patient 21 and 37 while the patient specific approach is superior for patient 36 and 39. The results indicate that if the patient have seizures which are dissimilar to normal non-seizure movements the patient specific approach is preferable. This is not difficult to grasp, if the seizures are outliers they can quite easily be identified using the patient specific approach. If the general approach is considered instead, the models will be trained to recognize subtle seizures as well. Since the subtle seizures are similar to non-seizure movements, it would result in more false positives, hence the patient specific approach gives better performance. For the patients with both representative and subtle seizures the general approach is prefered, this might be due to the increased seizure variability presented. In this case, the models are trained using more data which might be the reason as to why the general approach is superior.

6
Conclusion

Several machine learning algorithms have been implemented to consider a set of defined features and ANN has been used to evaluate pre-processed accelerometry directly. While most classifiers results in sufficient sensitivity this is usually at the cost of false positives. The results clearly indicate a trade-off which must be considered when using the classifiers in future work. If the current models are implemented in wearable electronics, it is up to the researcher in future work and physicians to decide how to handle this trade-off, whether the sensitivity or the specificity should be favored. KNN seems to be the most accurate method evaluated, and since it gives accurate result for GTCS classification as well it should be considered for multi-class classification. Additionally, the patient specific approach seems to be optimal for two patients while the general approach is superior in the remaining two. Hence there is no clear indication of which approach is optimal. As discussed, if seizures are removed due to a stricter seizure evaluation it is likely that the patient specific approach will be improved also for the cases where the general approach is currently favored.

Several methods have been considered to improve the specificity. The most efficient method was the implementation of veto filters in the post-processing layer. By removing seizure markings which does not fulfill certain conditions based on the distribution of seizure data in the training set the rate of false positives could be signficiantly reduced. Additonally, the amount of false positives could be reduced further by using the KDE method. However, two of the patiens had poor separability in the feature space which unfortunately resulted in a reduced amount of true positives as well.

The main difference between HMS and GTCS is the duration and activity. GTCS have an intense motor response which does not vary much among patients while HMS results in a wide range of motoric behavior. GTCS have a typically longer duration, between one and three minutes, while the duration of HMS is usually a couple of seconds up to one minute. By evaluating feature histograms it was verified that the distributions were different depending on type. Hence, the feature space in this thesis could serve as a basis for multi-class classification. Additionally, as suspected, the GTCS provided a smaller overlap with the distribution of normal data in the feature space, confirming that HMS is more difficult to detect than GTCS.

The current data set includes a subset of atypical or subtle seizures. Hence, given the data set available the results are not unexpected, the classifiers can detect HMS

with high sensitivity but it is difficult to do so with perfect accuracy. Some seizures are commonly overlooked due to their characteristics, they are either too subtle or too atypical to be detected using the remaining data. The subtle seizures are very similar to regular non-seizure movements, making the classifiers prone to give false alarms. The specificty is reduced for the patients with subtle or atypical seizures, which clearly indicates a correlation between atypical seizures and a larger amount false alarms. This indicates that the performance of both the general and the patient specific approach could be improved by removing the subtle and atypical seizures from the data sets. By entirely removing the seizures which are similar to non-seizure movements the feature separability would be improved and the classifiers would most likely achieve higher specificity. This is an issue which should be adressed by physicians, currently they only evaluate the video-EEG but ideally they should evaluate the accelerometer data as well to determine which seizures should be considered.

The limitations of the data set makes it unrealistic to expect perfect performance, there are simply too few seizures to represent the variability sufficiently. This especially affects the patient specific case, the classifiers are trained on a very small amount of seizures. Hence, more data should be recorded.

Bibliography

[1] A. Godfrey, R. Conway, D. Meagher, and G. ÓLaighin. Direct measurement of human movement by accelerometry. *Medical Engineering & Physics*, 30(10):1364 – 1386, 2008.

[2] Fisher. R. S. Epilepsy: A new definition. `http://www.epilepsy.com/article/2014/4/revised-definition-epilepsy`, 2014. Cited 2017-03.

[3] Sirven. J. I. and Shafer. P. O. What is epilepsy? `http://www.epilepsy.com/learn/epilepsy-101/what-epilepsy`, 2014. Cited 2017-03.

[4] A. Ulate-Campos, F. Coughlin, M. Gainza-Lein, I. S. Fernandez, P. L. Pearl, and T. Loddenkemper. Automated seizure detection systems and their effectiveness for each type of seizure. *Seizure*, 40:88–101, Aug 2016.

[5] Kiriakopoulos. E., Shafer. P. O., Fisher. R., and Sirven. J. I. Tonic-clonic seizures. `http://www.epilepsy.com/learn/types-seizures/tonic-clonic-seizures`, 2017. Cited 2017-03.

[6] K. Alqadi, R. Sankaraneni, U. Thome, and P. Kotagal. Semiology of hypermotor (hyperkinetic) seizures. *Epilepsy Behav*, 54:137–41, 2016.

[7] A. Van de Vel, K. Cuppens, B. Bonroy, M. Milosevic, S. Van Huffel, B. Vanrumste, L. Lagae, and B. Ceulemans. Long-term home monitoring of hypermotor seizures by patient-worn accelerometers. *Epilepsy Behav*, 26(1):118–25, 2013.

[8] K. Cuppens, P. Karsmakers, A. Van de Vel, B. Bonroy, M. Milosevic, S. Luca, T. Croonenborghs, B. Ceulemans, L. Lagae, S. Van Huffel, and B. Vanrumste. Accelerometry-based home monitoring for detection of nocturnal hypermotor seizures based on novelty detection. *IEEE J Biomed Health Inform*, 18(3):1026–33, 2014.

[9] M. Czarnecki and N. Gustafsson. Machine learning for detection of epileptic seizures. Master's thesis, Chalmers University of Technology, 2015.

[10] T. M. Nijsen, P. J. Cluitmans, J. B. Arends, and P. A. Griep. Detection of subtle nocturnal motor activity from 3-d accelerometry recordings in epilepsy patients. *IEEE Trans Biomed Eng*, 54(11):2073–81, 2007.

[11] MATLAB. *version 9.1.0.441655 (R2016b)*. The MathWorks Inc., Natick, Massachusetts, 2016.

[12] Hastie. T., Tibshirani. R., and Friedman. J. *The Elements of Statistical Learning, 2nd Edition*. Springer, online edition `http://statweb.stanford.edu/~tibs/ElemStatLearn/`, Stanford, California, 2008, Updated 2013.

[13] McBryan. M. Overlap integral. `https://chem.libretexts.org/Core/Physical_and_Theoretical_Chemistry/Quantum_Mechanics/11%3A_Molecules/Overlap_Integral`, 2017. Cited 2017-05-25.

[14] Zoubin Ghahramani. An introduction to hidden markov models and bayesian networks. *International Journal of Pattern Recognition and Artificial Intelligence*, 15(1):9–42, 2001.

A

Appendix 1: HMS Seizure data

Table A.1: HMS statistics describing seizure onsets, duration, pre-processed duration and standard deviation.

Patient	Start	Stop	Seizure time	Processed time	Max STD
21	2009-09-03 23:44:41	2009-09-03 23:44:54	00:13	00:13	0.5754
21	2009-09-04 01:21:18	2009-09-04 01:21:27	00:09	00:09	1.1323
21	2009-09-04 01:46:20	2009-09-04 01:46:37	00:17	00:17	0.1110
21	2009-09-04 07:22:02	2009-09-04 07:22:11	00:09	00:13	0.4729
21	2009-09-04 09:57:08	2009-09-04 09:57:18	00:10	00:10	0.3608
21	2009-09-04 12:49:09	2009-09-04 12:49:19	00:10	00:10	0.4948
21	2009-09-04 20:16:17	2009-09-04 20:16:35	00:18	00:04	0.3224
21	2009-09-05 03:41:41	2009-09-05 03:42:13	00:32	00:10	0.3782
36	2010-11-05 06:06:59	2010-11-05 06:07:30	00:31	00:31	0.6119
36	2010-11-07 01:54:28	2010-11-07 01:54:46	00:18	00:18	0.5953
36	2010-11-07 06:11:04	2010-11-07 06:11:35	00:31	00:31	0.6057
37	2011-02-09 22:55:38	2011-02-09 22:55:52	00:14	00:14	0.3660
37	2011-02-09 23:40:06	2011-02-09 23:40:25	00:19	00:19	0.5119
37	2011-02-10 01:02:46	2011-02-10 01:03:08	00:22	00:22	0.3222
37	2011-02-10 01:06:45	2011-02-10 01:06:56	00:11	00:11	0.5635
37	2011-02-10 01:14:52	2011-02-10 01:15:11	00:19	00:19	0.2864
37	2011-02-10 01:37:02	2011-02-10 01:37:14	00:12	00:12	0.2694
37	2011-02-10 01:46:07	2011-02-10 01:46:24	00:17	00:17	0.4035
37	2011-02-10 03:50:42	2011-02-10 03:50:52	00:10	00:10	0.3211
37	2011-02-10 04:55:43	2011-02-10 04:55:52	00:09	00:09	0.4991
37	2011-02-10 06:04:59	2011-02-10 06:05:12	00:13	00:13	0.4467
37	2011-02-10 08:48:24	2011-02-10 08:48:36	00:12	00:12	0.3016
37	2011-02-10 14:51:15	2011-02-10 14:51:22	00:07	00:07	0.2712
37	2011-02-10 22:30:39	2011-02-10 22:30:58	00:19	00:19	0.4002
37	2011-02-10 23:53:49	2011-02-10 23:54:05	00:16	00:15	0.4629
37	2011-02-11 00:21:56	2011-02-11 00:22:04	00:08	00:08	0.2183
37	2011-02-11 00:29:06	2011-02-11 00:29:17	00:11	00:06	0.1233
37	2011-02-11 00:34:05	2011-02-11 00:34:15	00:10	00:10	0.0759
37	2011-02-11 00:41:01	2011-02-11 00:41:12	00:11	00:11	0.2977
37	2011-02-11 02:01:03	2011-02-11 02:01:09	00:06	00:06	0.1938
37	2011-02-11 04:10:09	2011-02-11 04:10:24	00:15	00:15	0.3468
37	2011-02-11 04:21:17	2011-02-11 04:21:29	00:12	00:12	0.6351
37	2011-02-11 04:50:03	2011-02-11 04:50:18	00:15	00:15	0.3830
37	2011-02-11 06:09:15	2011-02-11 06:09:28	00:13	00:13	0.5977
37	2011-02-11 07:33:48	2011-02-11 07:34:09	00:21	00:21	0.5140
39	2011-03-24 17:36:53	2011-03-24 17:37:10	00:17	00:17	1.2042
39	2011-03-24 21:30:02	2011-03-24 21:30:15	00:13	00:12	0.9637

B

Appendix 2: HMS General result

B.1 Randomly balanced training set

Table B.1: General cross-validation results for KNN.

Method: KNN, Neighbors: 2, 5, 10								
KDE	Red		μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
		Accuracy	0.72755	0.75663	0.76227	0.091014	0.048143	0.060656
		Specificity	0.73868	0.76372	0.76969	0.098311	0.051383	0.065199
		Sensitivity	0.58616	0.60267	0.68733	0.13436	0.076772	0.09751
✗		Accuracy	0.89393	0.81721	0.86199	0.021867	0.042056	0.037568
		Specificity	0.90719	0.82537	0.87252	0.025079	0.045364	0.041591
		Sensitivity	0.3967	0.55627	0.51529	0.13083	0.10307	0.12703
	FS 2	Accuracy	0.72547	0.74073	0.75218	0.085418	0.062433	0.062941
		Specificity	0.73524	0.74572	0.7593	0.09226	0.065302	0.067786
		Sensitivity	0.65388	0.65208	0.66785	0.1349	0.092282	0.12991
	OI 2	Accuracy	0.71926	0.7111	0.70326	0.09635	0.070928	0.086018
		Specificity	0.72997	0.71415	0.70727	0.10278	0.073458	0.08936
		Sensitivity	0.58122	0.70823	0.72256	0.13467	0.074151	0.080707
✗	FS 2	Accuracy	0.87392	0.78223	0.81307	0.024567	0.044304	0.042218
		Specificity	0.88579	0.78877	0.82086	0.026568	0.047576	0.046264
		Sensitivity	0.39457	0.56823	0.59291	0.061385	0.082037	0.11734
✗	OI 2	Accuracy	0.86536	0.77733	0.81013	0.035282	0.059183	0.061652
		Specificity	0.87814	0.7831	0.81792	0.038099	0.061669	0.06443
		Sensitivity	0.39878	0.60877	0.58043	0.098039	0.08581	0.072209
	FS 5	Accuracy	0.7327	0.76382	0.76754	0.082177	0.043418	0.062812
		Specificity	0.74363	0.76992	0.77584	0.089233	0.046474	0.067771
		Sensitivity	0.59798	0.67486	0.66884	0.13805	0.11039	0.11965
	OI 5	Accuracy	0.72233	0.74054	0.73833	0.089475	0.05316	0.064645
		Specificity	0.73316	0.74472	0.74372	0.096144	0.055995	0.068404
		Sensitivity	0.58329	0.68895	0.6835	0.12093	0.089953	0.097069
✗	FS 5	Accuracy	0.87882	0.81135	0.84662	0.022962	0.042705	0.042888
		Specificity	0.89279	0.81777	0.85664	0.026143	0.044981	0.04677
		Sensitivity	0.31444	0.57997	0.52839	0.087905	0.087419	0.11106
✗	OI 5	Accuracy	0.87816	0.79413	0.83196	0.027456	0.049161	0.046029
		Specificity	0.89157	0.79967	0.83953	0.029945	0.051869	0.048804
		Sensitivity	0.33738	0.62863	0.56371	0.067604	0.10821	0.10038
	FS 10	Accuracy	0.72123	0.76461	0.77009	0.088047	0.047689	0.063586
		Specificity	0.73203	0.76961	0.77694	0.09499	0.050069	0.067441
		Sensitivity	0.56432	0.66968	0.68596	0.10403	0.080981	0.089564
	OI 10	Accuracy	0.73412	0.75849	0.76339	0.087496	0.048241	0.063199
		Specificity	0.74635	0.76432	0.77089	0.09478	0.051275	0.067322
		Sensitivity	0.57871	0.6522	0.65077	0.14672	0.10449	0.089
✗	FS 10	Accuracy	0.88712	0.8193	0.8557	0.020247	0.039331	0.042317
		Specificity	0.90034	0.82687	0.8657	0.023431	0.042324	0.046067
		Sensitivity	0.39119	0.55531	0.51598	0.11596	0.10441	0.1122
✗	OI 10	Accuracy	0.89056	0.81337	0.85553	0.022176	0.044248	0.039416
		Specificity	0.90493	0.82139	0.86575	0.02555	0.047624	0.043502
		Sensitivity	0.33593	0.53614	0.53407	0.097266	0.10422	0.12748

Table B.2: General cross-validation results for randf.

Method: Random Forest, Trees: 10, 30, 50								
KDE	Red		μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
		Accuracy	0.75149	0.78778	0.79032	0.065124	0.065081	0.06758
		Specificity	0.75472	0.79162	0.79477	0.067217	0.067556	0.070278
		Sensitivity	0.70074	0.69476	0.69228	0.079483	0.11332	0.11162
✗		Accuracy	0.80342	0.82928	0.83582	0.056924	0.056653	0.05391
		Specificity	0.80883	0.83628	0.84289	0.05911	0.059775	0.05704
		Sensitivity	0.59319	0.61036	0.59483	0.11061	0.12847	0.13149
	FS 2	Accuracy	0.71578	0.72993	0.72704	0.064269	0.063821	0.070806
		Specificity	0.71823	0.73277	0.73047	0.066533	0.066368	0.073499
		Sensitivity	0.69825	0.69728	0.71781	0.078278	0.11079	0.088166
	OI 2	Accuracy	0.70171	0.71237	0.71746	0.073604	0.072076	0.074094
		Specificity	0.70322	0.71491	0.72083	0.075581	0.074549	0.076538
		Sensitivity	0.73099	0.73868	0.68822	0.064715	0.09227	0.074052
✗	FS 2	Accuracy	0.79549	0.77838	0.79058	0.051726	0.06256	0.054313
		Specificity	0.80013	0.78283	0.79596	0.053809	0.064793	0.056651
		Sensitivity	0.62179	0.6301	0.60787	0.085793	0.10185	0.093289
✗	OI 2	Accuracy	0.76418	0.78323	0.78303	0.060517	0.05727	0.058179
		Specificity	0.76772	0.78901	0.78861	0.062323	0.05917	0.060213
		Sensitivity	0.66644	0.59091	0.60953	0.067507	0.055702	0.068473
	FS 5	Accuracy	0.73991	0.76109	0.75706	0.075628	0.067789	0.059222
		Specificity	0.74252	0.76515	0.76064	0.07814	0.070166	0.061565
		Sensitivity	0.72572	0.67467	0.71609	0.10469	0.093638	0.087911
	OI 5	Accuracy	0.72616	0.7409	0.74428	0.068422	0.066332	0.067667
		Specificity	0.72824	0.74366	0.74715	0.070684	0.068933	0.07018
		Sensitivity	0.73398	0.7324	0.73536	0.08186	0.1051	0.08784
✗	FS 5	Accuracy	0.79615	0.80013	0.81074	0.054694	0.060273	0.051169
		Specificity	0.80134	0.80693	0.81671	0.056955	0.063081	0.053542
		Sensitivity	0.62397	0.59193	0.61556	0.10047	0.091401	0.098161
✗	OI 5	Accuracy	0.77896	0.80009	0.80728	0.056415	0.057351	0.05564
		Specificity	0.78159	0.80606	0.81266	0.05844	0.060033	0.058134
		Sensitivity	0.70009	0.61763	0.63987	0.11073	0.1052	0.097319
	FS 10	Accuracy	0.74261	0.77025	0.76009	0.067012	0.063366	0.061855
		Specificity	0.74553	0.77427	0.76388	0.069432	0.065699	0.064421
		Sensitivity	0.72607	0.71341	0.73123	0.10824	0.096565	0.094299
	OI 10	Accuracy	0.74597	0.78076	0.78358	0.065539	0.064937	0.06271
		Specificity	0.74987	0.78597	0.78908	0.068306	0.067769	0.065452
		Sensitivity	0.70119	0.68918	0.66315	0.097117	0.10143	0.097131
✗	FS 10	Accuracy	0.80996	0.81325	0.8102	0.055136	0.054558	0.05632
		Specificity	0.81599	0.81965	0.81659	0.057512	0.056968	0.05941
		Sensitivity	0.6039	0.59218	0.63449	0.11429	0.065281	0.1256
✗	OI 10	Accuracy	0.80018	0.82751	0.82607	0.051245	0.050518	0.050048
		Specificity	0.80535	0.83433	0.8331	0.053709	0.053525	0.052774
		Sensitivity	0.61452	0.61313	0.59693	0.10062	0.098787	0.099239

Table B.3: General cross-validation results for logistic regression.

Method: Logistic Regression				
KDE	Red		μ_1	σ_1
		Accuracy	0.78148	0.059034
		Specificity	0.78788	0.062423
		Sensitivity	0.63573	0.12759
✗		Accuracy	0.83653	0.049902
		Specificity	0.84455	0.052536
		Sensitivity	0.52521	0.084945
	FS 2	Accuracy	0.74082	0.069243
		Specificity	0.74383	0.07216
		Sensitivity	0.71749	0.11301
	OI 2	Accuracy	0.72195	0.079889
		Specificity	0.72439	0.082558
		Sensitivity	0.74123	0.076078

✗	FS 2	Accuracy	0.7972	0.062538
		Specificity	0.80138	0.065318
		Sensitivity	0.68894	0.1181
✗	OI 2	Accuracy	0.80004	0.07551
		Specificity	0.80461	0.077775
		Sensitivity	0.63525	0.091716
	FS 5	Accuracy	0.75026	0.071022
		Specificity	0.75447	0.074352
		Sensitivity	0.70776	0.12339
	OI 5	Accuracy	0.73669	0.073058
		Specificity	0.73997	0.075785
		Sensitivity	0.70726	0.090904
✗	FS 5	Accuracy	0.81942	0.059748
		Specificity	0.82562	0.062843
		Sensitivity	0.64325	0.11544
✗	OI 5	Accuracy	0.8006	0.063983
		Specificity	0.80527	0.066343
		Sensitivity	0.65773	0.10157
	FS 10	Accuracy	0.77361	0.061375
		Specificity	0.77872	0.064538
		Sensitivity	0.66638	0.12843
	OI 10	Accuracy	0.76256	0.07007
		Specificity	0.76946	0.073948
		Sensitivity	0.63654	0.11616
✗	FS 10	Accuracy	0.82392	0.055779
		Specificity	0.83168	0.059066
		Sensitivity	0.55207	0.11363
✗	OI 10	Accuracy	0.81943	0.060406
		Specificity	0.82677	0.063418
		Sensitivity	0.56516	0.11346

Table B.4: General cross-validation results for SVM.

Method: SVM, Trees: 10, 30, 50										
KDE	Red		μ_1	μ_2	μ_3	μ_4	σ_1	σ_2	σ_3	σ_4
		Accuracy	0.57081	0.5706	0.79299	0.80163	0.11537	0.10729	0.060703	0.053203
		Specificity	0.58362	0.58268	0.80216	0.81015	0.13378	0.12424	0.065417	0.05704
		Sensitivity	0.70528	0.69607	0.60459	0.58765	0.22807	0.21873	0.12535	0.12847
✗		Accuracy	0.66973	0.66606	0.84995	0.85554	0.093235	0.091343	0.051934	0.042662
		Specificity	0.68137	0.67733	0.86018	0.86574	0.10687	0.10422	0.055013	0.045802
		Sensitivity	0.64627	0.64624	0.47039	0.47331	0.22017	0.21821	0.082485	0.077682
	FS 2	Accuracy	0.71773	0.70224	0.71924	0.71856	0.067354	0.053823	0.072883	0.072728
		Specificity	0.71958	0.70491	0.72095	0.72014	0.069776	0.056227	0.075856	0.075634
		Sensitivity	0.71394	0.70363	0.76157	0.76497	0.10149	0.085097	0.11109	0.10873
	OI 2	Accuracy	0.68499	0.70293	0.6931	0.69272	0.082281	0.07157	0.081858	0.081624
		Specificity	0.6853	0.70834	0.69358	0.69316	0.08445	0.073734	0.084468	0.084203
		Sensitivity	0.79574	0.58019	0.79763	0.79763	0.0656	0.055624	0.093219	0.093219
✗	FS 2	Accuracy	0.76945	0.75983	0.78283	0.79174	0.065753	0.053279	0.065842	0.06441
		Specificity	0.77584	0.76451	0.78645	0.79578	0.069951	0.055707	0.068938	0.067242
		Sensitivity	0.6626	0.66458	0.71961	0.69022	0.13436	0.11032	0.13404	0.12409
✗	OI 2	Accuracy	0.77985	0.76617	0.78717	0.7877	0.077756	0.061079	0.079578	0.079731
		Specificity	0.78521	0.77356	0.79102	0.79154	0.079953	0.063437	0.081838	0.081995
		Sensitivity	0.62699	0.51089	0.65594	0.65667	0.046848	0.049196	0.094414	0.094478
	FS 5	Accuracy	0.64646	0.63665	0.72619	0.74061	0.053995	0.050674	0.076854	0.070041
		Specificity	0.64463	0.63508	0.73005	0.74542	0.055647	0.052017	0.080979	0.073905
		Sensitivity	0.80166	0.78763	0.72531	0.67425	0.095046	0.070386	0.13113	0.1293
	OI 5	Accuracy	0.72664	0.70837	0.71767	0.71652	0.057549	0.051522	0.073929	0.074031
		Specificity	0.72978	0.7118	0.71949	0.71843	0.060205	0.053566	0.076522	0.076574
		Sensitivity	0.70933	0.66801	0.73812	0.72739	0.085992	0.08661	0.097831	0.096724
✗	FS 5	Accuracy	0.72709	0.70988	0.80682	0.83364	0.049415	0.044154	0.062601	0.062059
		Specificity	0.72804	0.71084	0.81357	0.84128	0.051185	0.045694	0.06644	0.065233
		Sensitivity	0.75567	0.72969	0.64158	0.58616	0.10876	0.083926	0.13054	0.119
		Accuracy	0.78905	0.77351	0.79434	0.79425	0.053618	0.043564	0.065801	0.065525

✗	OI 5	Specificity	0.79343	0.7783	0.79856	0.79865	0.055974	0.04574	0.068373	0.068161
		Sensitivity	0.65823	0.6105	0.67458	0.67164	0.11389	0.11054	0.11589	0.11616
	FS 10	Accuracy	0.61189	0.5836	0.76153	0.76009	0.092411	0.055826	0.06426	0.063924
		Specificity	0.62393	0.58013	0.76854	0.76563	0.10808	0.056896	0.067692	0.067135
		Sensitivity	0.6798	0.82951	0.59914	0.62548	0.20418	0.06645	0.090261	0.11361
	OI 10	Accuracy	0.5878	0.59145	0.74317	0.74929	0.053931	0.056821	0.076158	0.074792
		Specificity	0.58245	0.586	0.75081	0.75719	0.054338	0.057307	0.081651	0.080175
		Sensitivity	0.84958	0.84081	0.6553	0.64816	0.067944	0.055936	0.14137	0.14221
✗	FS 10	Accuracy	0.66831	0.67222	0.83951	0.84245	0.053323	0.046395	0.060459	0.059121
		Specificity	0.66863	0.67204	0.84729	0.85112	0.055473	0.048049	0.063504	0.062565
		Sensitivity	0.77144	0.76518	0.57891	0.54618	0.10013	0.09066	0.11953	0.13577
✗	OI 10	Accuracy	0.6878	0.68658	0.81691	0.81797	0.049511	0.051134	0.061645	0.060899
		Specificity	0.68686	0.68562	0.82526	0.82674	0.050859	0.052399	0.065575	0.064746
		Sensitivity	0.79495	0.79361	0.57256	0.54452	0.090106	0.077116	0.13314	0.12983

Table B.5: General post-processed results for KNN.

Method: KNN, Neighbors: 2, 5, 10										
			Original			Veto 1			Veto 2	
KDE	Red		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
		TP	35	36	36	27	32	31	25	27	28
		FP	432	451	404	237	282	274	170	195	192
		FN	2	1	1	10	5	6	12	10	9
✗		TP	18	25	22	11	17	12	9	14	7
		FP	244	351	288	110	149	126	67	106	87
		FN	19	12	15	26	20	25	28	23	30
	FS 2	TP	31	36	33	27	29	30	26	25	25
		FP	440	435	437	224	287	254	172	215	198
		FN	6	1	4	10	8	7	11	12	12
	OI 2	TP	33	36	36	29	32	32	24	29	28
		FP	394	432	421	239	277	266	200	208	212
		FN	4	1	1	8	5	5	13	8	9
✗	FS 2	TP	23	29	23	17	18	17	11	11	16
		FP	286	400	363	118	155	130	67	117	99
		FN	14	8	14	20	19	20	26	26	21
✗	OI 2	TP	22	27	26	17	19	15	10	16	13
		FP	286	370	331	122	138	133	82	119	106
		FN	15	10	11	20	18	22	27	21	24
	FS 5	TP	35	34	35	28	29	31	25	28	27
		FP	429	443	417	239	270	253	169	183	180
		FN	2	3	2	9	8	6	12	9	10
	OI 5	TP	35	35	36	29	30	32	26	27	27
		FP	414	457	427	220	281	270	184	200	192
		FN	2	2	1	8	7	5	11	10	10
✗	FS 5	TP	17	28	23	13	17	13	9	13	12
		FP	264	356	298	113	136	116	66	104	85
		FN	20	9	14	24	20	24	28	24	25
✗	OI 5	TP	23	27	24	16	17	14	11	12	10
		FP	262	365	322	107	145	112	64	109	88
		FN	14	10	13	21	20	23	26	25	27
	FS 10	TP	34	35	35	29	31	32	26	27	28
		FP	443	430	386	238	289	265	176	192	191
		FN	3	2	2	8	6	5	11	10	9
	OI 10	TP	33	37	36	28	33	33	23	26	28
		FP	414	428	398	227	272	263	177	197	179
		FN	4	0	1	9	4	4	14	11	9
✗	FS 10	TP	22	25	23	14	15	14	10	12	14
		FP	255	347	292	103	142	115	62	107	84
		FN	15	12	14	23	22	23	27	25	23
✗	OI 10	TP	21	26	20	15	17	11	10	11	9
		FP	248	337	286	105	132	115	69	103	90
		FN	16	11	17	22	20	26	27	26	28

Table B.6: General post-processed results for random forest.

Method: Random Forest, Trees: 10, 30, 50											
			Original			Veto 1			Veto 2		
KDE	Red		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
		TP	35	35	33	31	32	32	27	29	28
		FP	407	354	333	280	263	236	194	194	179
		FN	2	2	4	6	5	5	10	8	9
✗		TP	29	24	23	22	14	13	16	11	11
		FP	323	290	284	129	113	116	109	90	93
		FN	8	13	14	15	23	24	21	26	26
	FS 2	TP	35	35	36	32	30	33	28	27	24
		FP	446	438	414	270	280	277	212	206	210
		FN	2	2	1	5	7	4	9	10	13
	OI 2	TP	36	36	35	32	32	31	30	28	28
		FP	434	428	424	270	265	269	208	217	212
		FN	1	1	2	5	5	6	7	9	9
✗	FS 2	TP	28	28	27	17	20	17	12	13	14
		FP	360	365	357	132	140	127	116	117	106
		FN	9	9	10	20	17	20	25	24	23
✗	OI 2	TP	29	28	28	20	17	18	14	17	14
		FP	385	356	363	139	138	141	125	108	102
		FN	8	9	9	17	20	19	23	20	23
	FS 5	TP	34	34	35	32	32	30	29	27	26
		FP	416	404	389	270	279	283	193	183	200
		FN	3	3	2	5	5	7	8	10	11
	OI 5	TP	36	36	36	31	33	33	29	30	29
		FP	437	417	400	282	283	276	203	204	194
		FN	1	1	1	6	4	4	8	7	8
✗	FS 5	TP	28	27	28	18	17	17	16	13	15
		FP	342	341	335	131	121	125	113	106	97
		FN	9	10	9	19	20	20	21	24	22
✗	OI 5	TP	30	26	25	20	17	16	15	15	15
		FP	369	346	327	135	127	125	106	96	98
		FN	7	11	12	17	20	21	22	22	22
	FS 10	TP	34	36	35	29	32	32	27	28	27
		FP	413	392	381	259	274	265	192	187	204
		FN	3	1	2	8	5	5	10	9	10
	OI 10	TP	36	35	34	32	33	31	29	29	28
		FP	418	357	351	279	268	254	206	192	182
		FN	1	2	3	5	4	6	8	8	9
✗	FS 10	TP	30	26	24	21	14	17	15	14	14
		FP	331	339	331	130	129	123	104	102	105
		FN	7	11	13	16	23	20	22	23	23
✗	OI 10	TP	26	22	22	17	14	15	16	10	10
		FP	334	305	302	133	112	119	97	101	98
		FN	11	15	15	20	23	22	21	27	27

Table B.7: General post-processed results for logistic regression.

Method: Logistic Regression					
KDE	Red		Original	Veto 1	Veto 2
		TP	34	32	29
		FP	364	278	205
		FN	3	5	8
✗		TP	24	12	12
		FP	301	122	90
		FN	13	25	25
	FS 2	TP	34	31	28
		FP	393	271	215
		FN	3	6	9

	OI 2	TP	33	30	28
		FP	400	276	213
		FN	4	7	9
✗	FS 2	TP	26	16	13
		FP	313	119	98
		FN	11	21	24
✗	OI 2	TP	24	14	12
		FP	334	118	95
		FN	13	23	25
	FS 5	TP	34	30	27
		FP	373	273	212
		FN	3	7	10
	OI 5	TP	34	31	28
		FP	402	281	229
		FN	3	6	9
✗	FS 5	TP	24	13	10
		FP	296	123	95
		FN	13	24	27
✗	OI 5	TP	27	16	11
		FP	313	117	92
		FN	10	21	26
	FS 10	TP	33	30	27
		FP	361	259	197
		FN	4	7	10
	OI 10	TP	33	32	28
		FP	368	263	211
		FN	4	5	9
✗	FS 10	TP	24	13	13
		FP	302	122	96
		FN	13	24	24
✗	OI 10	TP	25	15	9
		FP	297	116	93
		FN	12	22	28

Table B.8: General post-processed results SVM.

Method: SVM, kernel: rbf, linear, BoxConst: 1, 100														
			Original				Veto 1				Veto 2			
KDE	Red		μ_1	μ_2	μ_3	μ_4	μ_1	μ_2	μ_3	μ_4	μ_1	μ_2	μ_3	μ_4
		TP	15	22	33	31	12	14	30	29	11	12	27	26
		FP	521	539	336	346	173	181	247	254	174	177	196	198
		FN	22	15	4	6	25	23	7	8	26	25	10	11
✗		TP	16	19	19	20	11	14	8	11	9	10	7	9
		FP	399	411	273	277	138	144	115	116	129	130	81	85
		FN	21	18	18	17	26	23	29	26	28	27	30	28
	FS 2	TP	35	34	33	34	30	29	31	31	30	30	28	28
		FP	449	506	416	417	304	280	280	280	223	215	224	223
		FN	2	3	4	3	7	8	6	6	7	7	9	9
	OI 2	TP	35	36	34	34	31	29	31	31	28	28	28	28
		FP	446	446	434	437	278	272	280	281	208	211	213	211
		FN	2	1	3	3	6	8	6	6	9	9	9	9
✗	FS 2	TP	24	32	25	24	14	21	15	15	13	17	13	13
		FP	365	422	318	313	138	145	127	121	109	113	100	96
		FN	13	5	12	13	23	16	22	22	24	20	24	24
✗	OI 2	TP	26	28	25	24	15	17	16	15	14	16	13	12
		FP	333	392	347	346	138	151	130	130	109	120	103	103
		FN	11	9	12	13	22	20	21	22	23	21	24	25
	FS 5	TP	34	36	34	34	31	31	30	30	29	27	29	28
		FP	581	573	408	398	287	269	269	271	222	218	215	219
		FN	3	1	3	3	6	6	7	7	8	10	8	9
	OI 5	TP	36	36	34	34	32	29	31	31	27	27	28	28
		FP	458	488	429	428	284	263	278	278	205	209	222	223
		FN	1	1	3	3	5	8	6	6	10	10	9	9

✗	FS 5	TP	31	34	21	19	20	23	12	12	18	16	10	9
		FP	457	463	300	286	159	174	124	113	112	127	100	85
		FN	6	3	16	18	17	14	25	25	19	21	27	28
✗	OI 5	TP	28	30	26	27	19	20	15	16	18	16	12	13
		FP	378	413	312	311	138	154	116	117	109	110	96	97
		FN	9	7	11	10	18	17	22	21	19	21	25	24
	FS 10	TP	22	37	34	33	18	30	30	30	14	29	25	27
		FP	537	640	374	393	229	272	260	261	193	220	212	197
		FN	15	0	3	4	19	7	7	7	23	8	12	10
	OI 10	TP	37	37	33	32	32	31	30	30	28	28	26	27
		FP	623	613	381	371	277	265	255	252	222	220	203	206
		FN	0	0	4	5	5	6	7	7	9	9	11	10
✗	FS 10	TP	33	33	19	22	21	21	12	14	19	18	10	9
		FP	486	493	273	284	158	166	112	112	140	138	84	79
		FN	4	4	18	15	16	16	25	23	18	19	27	28
✗	OI 10	TP	34	33	20	20	21	19	12	10	18	18	11	10
		FP	471	469	297	296	169	177	124	123	139	135	89	92
		FN	3	4	17	17	16	18	25	27	19	19	26	27

B.2 Cluster balanced training set

Table B.9: General cross-validation results for KNN.

Method: KNN, Neighbors: 2, 5, 10								
KDE	Red		μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
		Accuracy	0.80096	0.84639	0.82463	0.083063	0.012488	0.023367
		Specificity	0.82083	0.86267	0.84247	0.092997	0.01608	0.034169
		Sensitivity	0.36613	0.33948	0.35578	0.15782	0.072709	0.11104
✗		Accuracy	0.94852	0.84753	0.88151	0.0099548	0.021675	0.012163
		Specificity	0.96826	0.86055	0.89754	0.0049737	0.018842	0.011913
		Sensitivity	0.14906	0.31626	0.21314	0.058564	0.078103	0.026311
	FS 2	Accuracy	0.70218	0.46452	0.55874	0.046406	0.15016	0.11644
		Specificity	0.71122	0.45231	0.55893	0.052105	0.15915	0.1179
		Sensitivity	0.53175	0.64315	0.63341	0.095139	0.050418	0.02953
	OI 2	Accuracy	0.71415	0.74896	0.75032	0.091963	0.041419	0.056183
		Specificity	0.72798	0.75881	0.76336	0.098586	0.043448	0.061796
		Sensitivity	0.44511	0.4621	0.46033	0.11987	0.024792	0.071216
✗	FS 2	Accuracy	0.79542	0.46361	0.6062	0.038445	0.12627	0.090589
		Specificity	0.80667	0.45663	0.60107	0.039376	0.13065	0.097526
		Sensitivity	0.29011	0.58054	0.58694	0.035247	0.075599	0.066464
✗	OI 2	Accuracy	0.86461	0.75828	0.81169	0.021072	0.045101	0.033219
		Specificity	0.8805	0.76664	0.8235	0.023287	0.046955	0.036026
		Sensitivity	0.21162	0.44357	0.36902	0.026468	0.065054	0.055273
	FS 5	Accuracy	0.7094	0.72606	0.74716	0.053858	0.040369	0.03812
		Specificity	0.7223	0.73496	0.75812	0.062143	0.043066	0.043529
		Sensitivity	0.44377	0.4238	0.45668	0.07271	0.055255	0.042522
	OI 5	Accuracy	0.74213	0.7829	0.7814	0.097911	0.026635	0.038869
		Specificity	0.75775	0.79432	0.79517	0.10648	0.029913	0.045935
		Sensitivity	0.44576	0.43268	0.4325	0.13291	0.033088	0.046541
✗	FS 5	Accuracy	0.85689	0.76386	0.80526	0.028612	0.035421	0.030226
		Specificity	0.87121	0.7712	0.81639	0.029271	0.037637	0.031427
		Sensitivity	0.20701	0.47619	0.3187	0.044195	0.040618	0.016979
✗	OI 5	Accuracy	0.90479	0.79196	0.84705	0.0123	0.02904	0.018242
		Specificity	0.92189	0.80112	0.85976	0.012516	0.02914	0.020435
		Sensitivity	0.19381	0.42264	0.303	0.035864	0.039767	0.046483
	FS 10	Accuracy	0.77673	0.80544	0.79893	0.065864	0.022898	0.032222
		Specificity	0.79364	0.81944	0.81367	0.075724	0.026961	0.037208
		Sensitivity	0.42592	0.38956	0.35473	0.15765	0.065	0.063598

	OI 10	Accuracy	0.78387	0.83934	0.83713	0.088653	0.015542	0.026763
		Specificity	0.80248	0.8555	0.85564	0.097561	0.019775	0.037389
		Sensitivity	0.37331	0.34349	0.34112	0.12285	0.069769	0.094983
✗	FS 10	Accuracy	0.90906	0.81463	0.84963	0.020539	0.027316	0.017904
		Specificity	0.92698	0.8266	0.86254	0.022847	0.029122	0.016711
		Sensitivity	0.18335	0.35096	0.28655	0.04763	0.06106	0.035677
✗	OI 10	Accuracy	0.93322	0.84457	0.88933	0.009901	0.022241	0.011839
		Specificity	0.9522	0.85764	0.90548	0.0087529	0.020622	0.013814
		Sensitivity	0.15012	0.3121	0.23024	0.03712	0.070919	0.031781

Table B.10: General cross-validation results for randf.

Method: Random Forest, Trees: 10, 30, 50								
KDE	Red		μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
		Accuracy	0.39937	0.52311	0.47582	0.066256	0.10055	0.099686
		Specificity	0.38631	0.51659	0.46566	0.071972	0.10204	0.10441
		Sensitivity	0.74999	0.67425	0.67433	0.047927	0.050551	0.054626
✗		Accuracy	0.53975	0.58065	0.57238	0.061329	0.079141	0.080673
		Specificity	0.53303	0.57632	0.56631	0.064279	0.080459	0.083595
		Sensitivity	0.67172	0.61497	0.63527	0.087097	0.080131	0.076888
	FS 2	Accuracy	0.46296	0.57224	0.5638	0.10972	0.075957	0.060502
		Specificity	0.4498	0.57258	0.55983	0.12124	0.076547	0.0637
		Sensitivity	0.65498	0.51326	0.63436	0.077944	0.050529	0.059362
	OI 2	Accuracy	0.68573	0.71481	0.72507	0.051752	0.0457	0.047154
		Specificity	0.68926	0.72162	0.73212	0.052526	0.046707	0.048362
		Sensitivity	0.51976	0.47907	0.47542	0.043725	0.019992	0.029986
✗	FS 2	Accuracy	0.62888	0.73239	0.62009	0.06603	0.062199	0.0096489
		Specificity	0.62997	0.73923	0.62082	0.067577	0.063558	0.010997
		Sensitivity	0.51263	0.46302	0.50311	0.040514	0.044053	0.060066
✗	OI 2	Accuracy	0.697	0.71597	0.72306	0.059366	0.05812	0.057795
		Specificity	0.7033	0.72329	0.73148	0.060926	0.060177	0.059863
		Sensitivity	0.4571	0.46215	0.41295	0.05695	0.05791	0.071215
	FS 5	Accuracy	0.50046	0.55437	0.56055	0.094746	0.039592	0.023748
		Specificity	0.48957	0.55186	0.55745	0.10452	0.039995	0.02592
		Sensitivity	0.61503	0.60679	0.64688	0.086426	0.05382	0.093339
	OI 5	Accuracy	0.60021	0.65202	0.64443	0.049793	0.042594	0.040346
		Specificity	0.5983	0.65336	0.64439	0.051685	0.043321	0.041653
		Sensitivity	0.60818	0.59114	0.59887	0.035817	0.029131	0.037425
✗	FS 5	Accuracy	0.63478	0.59802	0.59489	0.041672	0.049977	0.031196
		Specificity	0.63598	0.59687	0.59511	0.042571	0.050434	0.031698
		Sensitivity	0.51981	0.55936	0.58314	0.02714	0.061073	0.071186
✗	OI 5	Accuracy	0.68199	0.69059	0.70891	0.057649	0.054776	0.052082
		Specificity	0.68441	0.69524	0.71334	0.0587	0.055889	0.05331
		Sensitivity	0.55015	0.49816	0.53886	0.056918	0.031757	0.039076
	FS 10	Accuracy	0.46942	0.56283	0.59341	0.10176	0.058653	0.036425
		Specificity	0.45575	0.5594	0.59071	0.11164	0.060546	0.037695
		Sensitivity	0.72781	0.60809	0.64068	0.062223	0.059292	0.042618
	OI 10	Accuracy	0.49824	0.54992	0.58094	0.086061	0.074265	0.070548
		Specificity	0.48988	0.54531	0.57795	0.090439	0.076682	0.071769
		Sensitivity	0.68398	0.6065	0.60268	0.04616	0.044992	0.046688
✗	FS 10	Accuracy	0.57509	0.62443	0.64378	0.046614	0.05887	0.041089
		Specificity	0.5683	0.62342	0.64428	0.052704	0.060204	0.04232
		Sensitivity	0.6509	0.58715	0.57234	0.091226	0.076221	0.070072
✗	OI 10	Accuracy	0.6068	0.63244	0.64399	0.084145	0.070491	0.065882
		Specificity	0.60619	0.63086	0.64311	0.085509	0.072139	0.068031
		Sensitivity	0.56162	0.57126	0.55502	0.051442	0.062001	0.054086

Table B.11: General cross-validation results for logistic regression.

Method: Logistic Regression

KDE	Red		μ_1	σ_1
		Accuracy	0.75653	0.034969
		Specificity	0.76231	0.037054
		Sensitivity	0.54502	0.075376
✗		Accuracy	0.7874	0.031942
		Specificity	0.79362	0.033473
		Sensitivity	0.49509	0.080091
	FS 2	Accuracy	0.64033	0.050336
		Specificity	0.64312	0.051447
		Sensitivity	0.53983	0.052516
	OI 2	Accuracy	0.78455	0.054939
		Specificity	0.79983	0.061071
		Sensitivity	0.32614	0.1101
✗	FS 2	Accuracy	0.75104	0.062937
		Specificity	0.76315	0.071205
		Sensitivity	0.41461	0.076192
✗	OI 2	Accuracy	0.66372	0.08624
		Specificity	0.677	0.093597
		Sensitivity	0.22102	0.050905
	FS 5	Accuracy	0.71935	0.063184
		Specificity	0.7297	0.069125
		Sensitivity	0.47496	0.067658
	OI 5	Accuracy	0.76699	0.054913
		Specificity	0.78014	0.060267
		Sensitivity	0.37338	0.11559
✗	FS 5	Accuracy	0.7454	0.045662
		Specificity	0.75537	0.051375
		Sensitivity	0.4119	0.026491
✗	OI 5	Accuracy	0.79782	0.044014
		Specificity	0.80959	0.047186
		Sensitivity	0.33033	0.12114
	FS 10	Accuracy	0.74259	0.041223
		Specificity	0.74823	0.043532
		Sensitivity	0.56194	0.072853
	OI 10	Accuracy	0.74309	0.049882
		Specificity	0.75577	0.055977
		Sensitivity	0.38386	0.082063
✗	FS 10	Accuracy	0.77497	0.036
		Specificity	0.78315	0.038063
		Sensitivity	0.43061	0.033705
✗	OI 10	Accuracy	0.77439	0.03927
		Specificity	0.78542	0.041867
		Sensitivity	0.33474	0.07438

Table B.12: General cross-validation results for SVM.

Method: SVM, Trees: 10, 30, 50										
KDE	Red		μ_1	μ_2	μ_3	μ_4	σ_1	σ_2	σ_3	σ_4
		Accuracy	0.94323	0.92415	0.75797	0.8085	0.012751	0.017713	0.039004	0.0018163
		Specificity	0.97011	0.94965	0.76615	0.81707	0.0063931	0.0064666	0.042875	0.0056199
		Sensitivity	0.025759	0.04911	0.50673	0.50821	0.019935	0.027982	0.063006	0.076296
✗		Accuracy	0.94331	0.93291	0.78359	0.84183	0.016753	0.016613	0.032976	0.017371
		Specificity	0.96492	0.95401	0.79081	0.851	0.007437	0.008927	0.035428	0.01744
		Sensitivity	0.021578	0.035008	0.49064	0.43741	0.013926	0.01687	0.098155	0.084837
	FS 2	Accuracy	0.76794	0.52901	0.97005	0.96866	0.033552	0.056822	0.015781	0.018167
		Specificity	0.78346	0.52769	0.99593	0.99259	0.038927	0.059269	0.0018502	0.0053274
		Sensitivity	0.25106	0.58387	0.070285	0.14156	0.074667	0.12396	0.022275	0.085939
	OI 2	Accuracy	0.71675	0.66635	0.78455	0.78386	0.058858	0.034413	0.062818	0.06281
		Specificity	0.72622	0.6714	0.80655	0.80588	0.061904	0.035672	0.076631	0.076706
		Sensitivity	0.43624	0.51135	0.26601	0.26335	0.054954	0.033317	0.1358	0.13554
✗	FS 2	Accuracy	0.71773	0.53217	0.97583	0.97611	0.058285	0.029594	0.012166	0.012704
		Specificity	0.72543	0.535	0.99694	0.9965	0.057045	0.03303	0.0016259	0.002048
		Sensitivity	0.26355	0.56834	0.0797	0.1389	0.071702	0.13812	0.030084	0.095037

✗	OI 2	Accuracy	0.75816	0.70393	0.89208	0.89178	0.047481	0.047784	0.081114	0.081414
		Specificity	0.76629	0.70939	0.91396	0.91366	0.049165	0.04877	0.086038	0.086344
		Sensitivity	0.38064	0.48781	0.032258	0.032258	0.091544	0.040213	0.032258	0.032258
	FS 5	Accuracy	0.86294	0.7795	0.73637	0.75526	0.010937	0.031043	0.065791	0.069067
		Specificity	0.88548	0.79486	0.75111	0.7712	0.019835	0.024814	0.076676	0.078779
		Sensitivity	0.13882	0.22183	0.45291	0.39033	0.050974	0.037261	0.081431	0.064579
	OI 5	Accuracy	0.76451	0.72715	0.77055	0.76035	0.032831	0.032234	0.065506	0.056189
		Specificity	0.77318	0.73327	0.79187	0.77391	0.035463	0.03363	0.080109	0.063249
		Sensitivity	0.52108	0.5762	0.2884	0.37827	0.05405	0.080685	0.16213	0.13243
✗	FS 5	Accuracy	0.84457	0.81772	0.77742	0.80112	0.014311	0.026874	0.042608	0.052581
		Specificity	0.86131	0.83199	0.79043	0.81493	0.018223	0.024155	0.051467	0.06123
		Sensitivity	0.15032	0.18427	0.38712	0.39023	0.043276	0.041773	0.07141	0.082163
✗	OI 5	Accuracy	0.77074	0.75646	0.80969	0.81003	0.032285	0.031914	0.054405	0.056815
		Specificity	0.77808	0.76319	0.82775	0.8276	0.032338	0.033213	0.065654	0.066994
		Sensitivity	0.46784	0.53006	0.254	0.25908	0.034111	0.084195	0.15219	0.14504
	FS 10	Accuracy	0.91515	0.86346	0.75287	0.74909	0.0082403	0.023733	0.052549	0.03514
		Specificity	0.94055	0.88519	0.76604	0.75903	0.012023	0.016453	0.0607	0.038661
		Sensitivity	0.063072	0.10079	0.41692	0.40314	0.040153	0.03965	0.049819	0.032565
	OI 10	Accuracy	0.89868	0.87098	0.7238	0.7298	0.012496	0.01564	0.060147	0.062101
		Specificity	0.92318	0.89357	0.73834	0.7452	0.018744	0.012598	0.070009	0.07179
		Sensitivity	0.090815	0.10313	0.37767	0.34325	0.040022	0.041705	0.09556	0.089901
✗	FS 10	Accuracy	0.92235	0.8855	0.76844	0.78086	0.011256	0.02173	0.041695	0.037603
		Specificity	0.94297	0.90386	0.77818	0.79141	0.0092369	0.01627	0.046152	0.043061
		Sensitivity	0.060333	0.093093	0.46936	0.44222	0.030192	0.034148	0.086476	0.055236
✗	OI 10	Accuracy	0.90853	0.89511	0.74773	0.75012	0.012802	0.016262	0.048585	0.041178
		Specificity	0.92863	0.91428	0.75959	0.76153	0.010452	0.01276	0.054073	0.04514
		Sensitivity	0.068608	0.078148	0.33879	0.32892	0.029362	0.031052	0.084611	0.072033

Table B.13: General post-processed results for KNN.

Method: KNN, Neighbors: 2, 5, 10											
			Original			Veto 1			Veto 2		
KDE	Red		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
		TP	19	32	27	15	23	21	9	11	12
		FP	285	380	384	106	205	179	85	86	103
		FN	18	5	10	22	14	16	28	26	25
✗		TP	3	25	21	2	18	16	2	10	8
		FP	94	351	290	35	142	119	8	86	59
		FN	34	12	16	35	19	21	35	27	29
	FS 2	TP	36	37	36	25	20	31	22	27	27
		FP	539	570	562	231	215	182	181	199	210
		FN	1	0	1	12	17	6	15	10	10
	OI 2	TP	31	36	34	22	29	27	19	24	20
		FP	428	457	441	213	229	245	164	186	175
		FN	6	1	3	15	8	10	18	13	17
✗	FS 2	TP	30	33	34	21	21	20	19	24	23
		FP	428	478	477	184	140	157	125	185	149
		FN	7	4	3	16	16	17	18	13	14
✗	OI 2	TP	25	29	29	15	21	22	8	16	13
		FP	314	422	386	127	163	158	69	103	101
		FN	12	8	8	22	16	15	29	21	24
	FS 5	TP	35	36	36	25	29	29	13	20	19
		FP	538	495	480	226	221	212	174	183	161
		FN	2	1	1	12	8	8	24	17	18
	OI 5	TP	30	33	33	23	25	24	18	21	17
		FP	371	442	420	158	218	208	125	141	141
		FN	7	4	4	14	12	13	19	16	20
✗	FS 5	TP	26	29	28	16	21	18	10	13	12
		FP	317	412	394	149	159	151	81	107	97
		FN	11	8	9	21	16	19	27	24	25
✗	OI 5	TP	19	29	24	13	18	14	6	18	8
		FP	231	408	355	95	167	130	44	107	81
		FN	18	8	13	24	19	23	31	19	29

	FS 10	TP	29	35	33	23	26	25	14	13	19
		FP	385	433	426	169	210	220	113	126	129
		FN	8	2	4	14	11	12	23	24	18
	OI 10	TP	26	30	25	20	22	17	12	14	11
		FP	323	375	364	129	189	153	109	90	88
		FN	11	7	12	17	15	20	25	23	26
✗	FS 10	TP	12	26	28	7	18	19	5	10	12
		FP	199	393	350	88	155	140	42	98	79
		FN	25	11	9	30	19	18	32	27	25
✗	OI 10	TP	12	25	16	7	18	7	5	11	5
		FP	141	344	268	45	139	99	22	75	54
		FN	25	12	21	30	19	30	32	26	32

Table B.14: General post-processed results for random forest.

Method: Random Forest, Trees: 10, 30, 50											
			Original			Veto 1			Veto 2		
KDE	Red		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
		TP	36	36	36	14	21	19	25	26	26
		FP	683	588	611	124	164	154	209	205	204
		FN	1	1	1	23	16	18	12	11	11
✗		TP	35	35	35	24	24	22	26	27	27
		FP	510	483	467	165	148	150	208	204	203
		FN	2	2	2	13	13	15	11	10	10
	FS 2	TP	37	36	36	27	25	18	27	22	25
		FP	663	579	589	262	173	170	185	196	184
		FN	0	1	1	10	12	19	10	15	12
	OI 2	TP	36	36	36	27	26	27	25	26	26
		FP	490	459	447	226	236	238	189	184	206
		FN	1	1	1	10	11	10	12	11	11
✗	FS 2	TP	35	35	35	23	26	21	27	23	23
		FP	466	402	497	185	201	142	194	173	174
		FN	2	2	2	14	11	16	10	14	14
✗	OI 2	TP	35	35	35	25	29	29	27	24	25
		FP	418	412	395	185	186	194	182	172	167
		FN	2	2	2	12	8	8	10	13	12
	FS 5	TP	36	36	36	16	21	20	27	25	24
		FP	628	604	589	166	145	137	199	199	189
		FN	1	1	1	21	16	17	10	12	13
	OI 5	TP	36	36	36	22	21	23	22	22	23
		FP	577	516	519	169	190	161	188	197	190
		FN	1	1	1	15	16	14	15	15	14
✗	FS 5	TP	35	35	35	21	21	23	23	25	25
		FP	496	501	494	170	135	134	190	188	182
		FN	2	2	2	16	16	14	14	12	12
✗	OI 5	TP	35	35	35	26	23	24	23	22	23
		FP	426	417	407	185	168	177	185	179	172
		FN	2	2	2	11	14	13	14	15	14
	FS 10	TP	36	36	35	21	20	22	26	24	25
		FP	618	588	544	197	143	153	206	187	189
		FN	1	1	2	16	17	15	11	13	12
	OI 10	TP	36	36	36	20	25	23	25	27	27
		FP	605	582	556	147	173	173	200	208	199
		FN	1	1	1	17	12	14	12	10	10
✗	FS 10	TP	35	35	35	22	23	22	26	25	24
		FP	502	459	457	151	159	151	197	190	186
		FN	2	2	2	15	14	15	11	12	13
✗	OI 10	TP	35	35	35	26	22	25	26	26	26
		FP	440	437	441	168	165	182	202	196	200
		FN	2	2	2	11	15	12	11	11	11

Table B.15: General post-processed results for logistic regression.

Method: Logistic Regression					
KDE	Red		Original	Veto 1	Veto 2
		TP	35	30	25
		FP	447	259	181
		FN	2	7	12
✗		TP	30	21	19
		FP	385	148	108
		FN	7	16	18
	FS 2	TP	36	25	27
		FP	531	197	210
		FN	1	12	10
	OI 2	TP	27	21	17
		FP	385	223	165
		FN	10	16	20
✗	FS 2	TP	17	10	8
		FP	358	125	103
		FN	20	27	29
✗	OI 2	TP	26	13	13
		FP	425	107	109
		FN	11	24	24
	FS 5	TP	34	27	24
		FP	486	192	192
		FN	3	10	13
	OI 5	TP	32	28	21
		FP	409	253	186
		FN	5	9	16
✗	FS 5	TP	28	19	15
		FP	399	132	108
		FN	9	18	22
✗	OI 5	TP	26	18	12
		FP	362	128	99
		FN	11	19	25
	FS 10	TP	36	29	25
		FP	462	250	193
		FN	1	8	12
	OI 10	TP	31	30	24
		FP	444	245	198
		FN	6	7	13
✗	FS 10	TP	31	23	18
		FP	405	145	114
		FN	6	14	19
✗	OI 10	TP	28	20	16
		FP	392	135	112
		FN	9	17	21

Table B.16: General post-processed results SVM.

Method: SVM, kernel: rbf, linear, BoxConst: 1, 100														
			Original				Veto 1				Veto 2			
KDE	Red		μ_1	μ_2	μ_3	μ_4	μ_1	μ_2	μ_3	μ_4	μ_1	μ_2	μ_3	μ_4
		TP	10	14	34	35	1	2	29	31	0	0	24	25
		FP	112	183	430	424	32	39	243	252	0	0	166	161
		FN	27	23	3	2	36	35	8	6	37	37	13	12
✗		TP	9	10	30	29	6	7	21	19	1	3	15	13
		FP	117	150	384	351	39	43	139	126	11	16	110	79
		FN	28	27	7	8	31	30	16	18	36	34	22	24
	FS 2	TP	31	37	8	12	25	20	8	11	15	23	4	10
		FP	430	639	14	28	235	177	9	18	150	189	5	10
		FN	6	0	29	25	12	17	29	26	22	14	33	27
	OI 2	TP	36	36	11	11	28	25	7	7	25	26	5	5
		FP	459	562	348	349	231	207	187	187	187	202	140	139

		FN	1	1	26	26	9	12	30	30	12	11	32	32
✗	FS 2	TP	31	35	2	3	17	21	1	1	20	26	0	0
		FP	439	540	12	13	143	146	4	5	125	178	0	0
		FN	6	2	35	34	20	16	36	36	17	11	37	37
✗	OI 2	TP	29	33	2	2	19	26	2	2	17	18	2	2
		FP	408	459	79	79	148	179	45	46	122	133	43	43
		FN	8	4	35	35	18	11	35	35	20	19	35	35
	FS 5	TP	16	33	26	28	8	18	23	24	4	9	19	20
		FP	350	483	434	403	161	172	214	189	65	89	170	180
		FN	21	4	11	9	29	19	14	13	33	28	18	17
	OI 5	TP	36	35	11	31	30	27	8	26	25	24	6	21
		FP	452	506	367	431	240	248	196	237	168	174	147	173
		FN	1	2	26	6	7	10	29	11	12	13	31	16
✗	FS 5	TP	24	26	22	16	15	21	14	10	9	16	11	7
		FP	360	389	398	317	132	154	139	105	77	88	107	99
		FN	13	11	15	21	22	16	23	27	28	21	26	30
✗	OI 5	TP	29	28	8	9	19	21	6	7	18	11	4	4
		FP	407	435	316	313	152	167	116	120	113	123	92	92
		FN	8	9	29	28	18	16	31	30	19	26	33	33
	FS 10	TP	11	24	29	33	4	10	23	28	0	2	22	22
		FP	204	348	408	465	82	119	231	264	11	30	169	188
		FN	26	13	8	4	33	27	14	9	37	35	15	15
	OI 10	TP	17	24	28	29	4	9	24	26	1	1	18	19
		FP	245	341	431	427	101	119	205	200	23	38	178	175
		FN	20	13	9	8	33	28	13	11	36	36	19	18
✗	FS 10	TP	12	19	29	27	9	15	21	20	3	9	11	12
		FP	182	289	386	379	68	116	134	131	29	41	115	106
		FN	25	18	8	10	28	22	16	17	34	28	26	25
✗	OI 10	TP	12	17	26	26	7	13	17	18	4	7	10	13
		FP	220	238	393	408	73	85	132	136	29	47	111	107
		FN	25	20	11	11	30	24	20	19	33	30	27	24

C

Appendix 3: HMS Patient 21 result

C.1 Randomly balanced training set

Table C.1: Cross-validation results for KNN, patient 21.

Method: KNN, Neighbors: 2, 5, 10								
KDE	Red		μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
		Accuracy	0.89307	0.82861	0.88911	0.0094919	0.015675	0.0099372
		Specificity	0.90111	0.83279	0.89521	0.0099704	0.015626	0.0095423
		Sensitivity	0.29425	0.5421	0.46932	0.054593	0.063122	0.074267
✗		Accuracy	0.93239	0.88358	0.93625	0.0079124	0.013007	0.0088568
		Specificity	0.93926	0.88879	0.94283	0.0083238	0.013303	0.0090089
		Sensitivity	0.24795	0.38353	0.30905	0.066477	0.082433	0.069643
	FS 2	Accuracy	0.88628	0.81583	0.86163	0.010582	0.015532	0.021091
		Specificity	0.8935	0.8195	0.86657	0.0097642	0.015465	0.020805
		Sensitivity	0.33844	0.56645	0.49607	0.062795	0.065513	0.093553
	OI 2	Accuracy	0.87723	0.81125	0.86591	0.010392	0.016011	0.014476
		Specificity	0.88362	0.81437	0.87074	0.010053	0.015633	0.013717
		Sensitivity	0.37545	0.58447	0.5162	0.089197	0.084839	0.083102
✗	FS 2	Accuracy	0.92503	0.87902	0.93135	0.0054679	0.012404	0.010736
		Specificity	0.9312	0.88381	0.93654	0.0053946	0.012975	0.0098928
		Sensitivity	0.35234	0.49308	0.46769	0.056864	0.077618	0.091859
✗	OI 2	Accuracy	0.94077	0.8853	0.92224	0.008624	0.016917	0.013105
		Specificity	0.94779	0.89013	0.92772	0.0085196	0.017154	0.012827
		Sensitivity	0.24503	0.45493	0.4211	0.062112	0.089181	0.083241
	FS 5	Accuracy	0.87826	0.81023	0.87661	0.007275	0.018888	0.014137
		Specificity	0.88482	0.81433	0.88113	0.0072223	0.018612	0.013233
		Sensitivity	0.39217	0.51863	0.53467	0.084223	0.077206	0.10026
	OI 5	Accuracy	0.88377	0.82034	0.89098	0.010102	0.011305	0.012408
		Specificity	0.89096	0.82459	0.89683	0.0097943	0.010995	0.011837
		Sensitivity	0.35045	0.51665	0.44863	0.072152	0.069791	0.076622
✗	FS 5	Accuracy	0.92073	0.88053	0.94318	0.0050792	0.015343	0.0075848
		Specificity	0.92756	0.88464	0.94918	0.0048617	0.015855	0.0063359
		Sensitivity	0.27662	0.53883	0.36175	0.054106	0.078871	0.10216
✗	OI 5	Accuracy	0.93507	0.88372	0.93653	0.0062346	0.010217	0.0084868
		Specificity	0.94213	0.88835	0.94249	0.0060525	0.010589	0.0076667
		Sensitivity	0.2348	0.47107	0.39058	0.066624	0.080279	0.067711
	FS 10	Accuracy	0.88622	0.81056	0.87796	0.011532	0.017727	0.012542
		Specificity	0.89289	0.81438	0.88254	0.011403	0.017804	0.012628
		Sensitivity	0.38662	0.53054	0.55431	0.068337	0.06776	0.080013
	OI 10	Accuracy	0.88883	0.8145	0.88122	0.0087497	0.015878	0.010891
		Specificity	0.89579	0.81819	0.88623	0.0092358	0.015951	0.010289
		Sensitivity	0.37578	0.55164	0.52594	0.070513	0.067409	0.082857
✗	FS 10	Accuracy	0.92705	0.88088	0.9353	0.0073375	0.010764	0.0079205
		Specificity	0.93357	0.88559	0.94106	0.0075395	0.010941	0.0074027
		Sensitivity	0.28318	0.4524	0.38964	0.074602	0.08657	0.081072
✗	OI 10	Accuracy	0.94066	0.88811	0.94177	0.0054291	0.0084594	0.0057679
		Specificity	0.94747	0.89291	0.94738	0.0049563	0.0083771	0.0042887

		Sensitivity	0.26532	0.44775	0.41425	0.05822	0.071481	0.077403

Table C.2: Cross-validation results for randf, patient 21.

Method: Random Forest, Trees: 10, 30, 50								
KDE	Red		μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
		Accuracy	0.81098	0.84163	0.85497	0.020298	0.019961	0.017474
		Specificity	0.81412	0.84538	0.8589	0.020605	0.019502	0.017116
		Sensitivity	0.58171	0.57148	0.55785	0.097639	0.073278	0.089082
✗		Accuracy	0.87097	0.91738	0.91632	0.01072	0.004355	0.0078249
		Specificity	0.87472	0.92205	0.92133	0.011682	0.0040029	0.0078532
		Sensitivity	0.56886	0.52078	0.4676	0.10668	0.10222	0.083972
	FS 2	Accuracy	0.79315	0.81515	0.81731	0.026073	0.018813	0.016797
		Specificity	0.79589	0.81834	0.82093	0.026463	0.018716	0.016456
		Sensitivity	0.57976	0.56199	0.53431	0.093119	0.076994	0.086364
	OI 2	Accuracy	0.78683	0.80398	0.81195	0.01991	0.025055	0.025922
		Specificity	0.78881	0.80695	0.81491	0.02008	0.024739	0.025731
		Sensitivity	0.63452	0.5869	0.59827	0.075304	0.076138	0.082401
✗	FS 2	Accuracy	0.85538	0.89402	0.90222	0.015845	0.010149	0.012162
		Specificity	0.85997	0.89814	0.90681	0.015819	0.0093501	0.010919
		Sensitivity	0.42477	0.53363	0.49935	0.074121	0.093574	0.10021
✗	OI 2	Accuracy	0.86192	0.88805	0.89391	0.016623	0.014997	0.013999
		Specificity	0.86608	0.89301	0.89904	0.016823	0.015099	0.013875
		Sensitivity	0.50915	0.44844	0.42484	0.09784	0.089942	0.099888
	FS 5	Accuracy	0.80106	0.82954	0.83526	0.02031	0.015666	0.020227
		Specificity	0.80419	0.83271	0.83894	0.020247	0.015095	0.020267
		Sensitivity	0.5619	0.59713	0.54612	0.091364	0.08641	0.077663
	OI 5	Accuracy	0.78629	0.81534	0.82617	0.021776	0.022264	0.023459
		Specificity	0.78825	0.81861	0.82973	0.022131	0.022328	0.023993
		Sensitivity	0.64015	0.56677	0.55586	0.10618	0.099797	0.10148
✗	FS 5	Accuracy	0.86622	0.90288	0.90311	0.013028	0.0091133	0.010059
		Specificity	0.87026	0.90745	0.90757	0.013646	0.008581	0.0093759
		Sensitivity	0.50044	0.49493	0.50873	0.072663	0.085613	0.09233
✗	OI 5	Accuracy	0.8545	0.89833	0.90316	0.017716	0.0131	0.012722
		Specificity	0.85806	0.90282	0.90784	0.017876	0.013405	0.012634
		Sensitivity	0.56295	0.51175	0.49935	0.10243	0.095442	0.09413
	FS 10	Accuracy	0.80274	0.83125	0.83497	0.028426	0.017769	0.020126
		Specificity	0.80559	0.83519	0.83824	0.028321	0.017616	0.01993
		Sensitivity	0.59895	0.52781	0.59583	0.089815	0.087199	0.084853
	OI 10	Accuracy	0.78631	0.82545	0.84526	0.021807	0.023146	0.018494
		Specificity	0.78858	0.82829	0.84837	0.022075	0.023367	0.018093
		Sensitivity	0.62551	0.62145	0.61414	0.084448	0.08738	0.079623
✗	FS 10	Accuracy	0.86565	0.90229	0.90805	0.012739	0.010361	0.01682
		Specificity	0.87034	0.90686	0.91277	0.013027	0.0097121	0.016481
		Sensitivity	0.47182	0.50828	0.49805	0.074887	0.096469	0.097755
✗	OI 10	Accuracy	0.85751	0.9014	0.90561	0.0183	0.012137	0.011437
		Specificity	0.86149	0.90622	0.91057	0.018527	0.01231	0.011075
		Sensitivity	0.5313	0.47717	0.47013	0.087852	0.090564	0.09148

Table C.3: Cross-validation results for logistic regression, patient 21.

Method: Logistic Regression				
KDE	Red		μ_1	σ_1
		Accuracy	0.75014	0.010641
		Specificity	0.75294	0.011299
		Sensitivity	0.60132	0.098643
✗		Accuracy	0.83388	0.012016
		Specificity	0.83871	0.012998
		Sensitivity	0.43613	0.081429
	FS 2	Accuracy	0.84253	0.013759
		Specificity	0.84742	0.013231
		Sensitivity	0.48994	0.114

	OI 2	Accuracy	0.83746	0.017044
		Specificity	0.84078	0.016946
		Sensitivity	0.58905	0.10744
✗	FS 2	Accuracy	0.92323	0.011612
		Specificity	0.92856	0.011352
		Sensitivity	0.45097	0.089625
✗	OI 2	Accuracy	0.92135	0.010128
		Specificity	0.926	0.0094545
		Sensitivity	0.52321	0.10477
	FS 5	Accuracy	0.81256	0.016968
		Specificity	0.81663	0.016957
		Sensitivity	0.51313	0.063961
	OI 5	Accuracy	0.83135	0.017224
		Specificity	0.83445	0.017194
		Sensitivity	0.61077	0.085496
✗	FS 5	Accuracy	0.90631	0.011256
		Specificity	0.91138	0.010775
		Sensitivity	0.4724	0.08517
✗	OI 5	Accuracy	0.91605	0.010185
		Specificity	0.92078	0.0088446
		Sensitivity	0.49389	0.096334
	FS 10	Accuracy	0.78525	0.011862
		Specificity	0.78945	0.012222
		Sensitivity	0.47339	0.077983
	OI 10	Accuracy	0.82046	0.012927
		Specificity	0.82413	0.012628
		Sensitivity	0.59737	0.087889
✗	FS 10	Accuracy	0.87315	0.01059
		Specificity	0.87792	0.0093055
		Sensitivity	0.47655	0.10991
✗	OI 10	Accuracy	0.89866	0.0077734
		Specificity	0.90376	0.0069257
		Sensitivity	0.46347	0.07986

Table C.4: Cross-validation results for SVM, patient 21.

Method: SVM, Trees: 10, 30, 50										
KDE	Red		μ_1	μ_2	μ_3	μ_4	σ_1	σ_2	σ_3	σ_4
		Accuracy	0.96465	0.94504	0.82982	0.75277	0.0038707	0.0056756	0.012263	0.015988
		Specificity	0.97727	0.95729	0.83446	0.75646	0.0045027	0.0067644	0.012566	0.016959
		Sensitivity	0.011364	0.020292	0.51843	0.53009	0.011364	0.013411	0.075691	0.082138
✗		Accuracy	0.98464	0.96257	0.90434	0.82786	0.0031454	0.0044492	0.0069245	0.015503
		Specificity	0.99399	0.97148	0.91033	0.83252	0.0025536	0.0051884	0.0061002	0.01588
		Sensitivity	0.012401	0.03373	0.3636	0.44808	0.0091059	0.026694	0.058215	0.081605
	FS 2	Accuracy	0.76617	0.76175	0.86864	0.85221	0.023744	0.028746	0.015659	0.01387
		Specificity	0.76839	0.7676	0.87311	0.85703	0.023817	0.028721	0.01538	0.014604
		Sensitivity	0.61008	0.29697	0.54594	0.50779	0.071437	0.070475	0.10976	0.10574
	OI 2	Accuracy	0.8347	0.74431	0.84583	0.84333	0.015391	0.016024	0.011498	0.011267
		Specificity	0.83836	0.74671	0.84926	0.84674	0.014448	0.016209	0.010945	0.010797
		Sensitivity	0.56418	0.58277	0.58905	0.58905	0.072249	0.0578	0.10744	0.10744
✗	FS 2	Accuracy	0.85541	0.85555	0.92856	0.92979	0.0081788	0.019447	0.010292	0.0119
		Specificity	0.85886	0.86246	0.93388	0.9351	0.0084988	0.01914	0.0096119	0.011048
		Sensitivity	0.5873	0.15068	0.44627	0.45932	0.090098	0.040065	0.089543	0.083607
✗	OI 2	Accuracy	0.92037	0.85226	0.93261	0.92891	0.014042	0.021137	0.009761	0.010105
		Specificity	0.92565	0.8567	0.93782	0.93398	0.013509	0.021207	0.0087481	0.0089884
		Sensitivity	0.4474	0.47418	0.46526	0.47419	0.09163	0.058567	0.085991	0.088006
	FS 5	Accuracy	0.82024	0.78806	0.84229	0.82422	0.045971	0.033746	0.011091	0.013427
		Specificity	0.8276	0.79365	0.84706	0.82947	0.047167	0.034456	0.010679	0.013867
		Sensitivity	0.22693	0.33829	0.5138	0.45301	0.10637	0.071308	0.098583	0.070655
	OI 5	Accuracy	0.83404	0.78154	0.83988	0.83318	0.0202	0.013641	0.013479	0.014966
		Specificity	0.83989	0.78646	0.84325	0.83616	0.020407	0.013667	0.013214	0.014627
		Sensitivity	0.39352	0.40112	0.60467	0.61658	0.079562	0.056793	0.09983	0.082477
		Accuracy	0.90871	0.85779	0.92939	0.90986	0.024277	0.014876	0.008152	0.010305

✗	FS 5	Specificity	0.91624	0.86439	0.93506	0.91637	0.023024	0.013877	0.0076373	0.0098303
		Sensitivity	0.091585	0.17915	0.41792	0.29503	0.046884	0.040801	0.067519	0.053455
✗	OI 5	Accuracy	0.90131	0.85586	0.93106	0.92753	0.018183	0.014252	0.0089971	0.010219
		Specificity	0.90816	0.86187	0.93614	0.93264	0.016961	0.014204	0.0075762	0.0085757
		Sensitivity	0.18779	0.25656	0.47662	0.46769	0.054503	0.070835	0.094682	0.088978
	FS 10	Accuracy	0.89991	0.89623	0.8296	0.77989	0.044974	0.011623	0.014225	0.010251
		Specificity	0.91034	0.9074	0.83439	0.78299	0.045713	0.012196	0.014828	0.0098547
		Sensitivity	0.061959	0.045094	0.52603	0.58297	0.040919	0.019177	0.095168	0.091103
	OI 10	Accuracy	0.94127	0.90234	0.8443	0.82599	0.0045649	0.0063983	0.014228	0.016024
		Specificity	0.95331	0.91389	0.84798	0.83014	0.0052101	0.0073706	0.014017	0.015905
		Sensitivity	0.029221	0.029221	0.61851	0.57127	0.019749	0.019749	0.098064	0.096858
✗	FS 10	Accuracy	0.97305	0.91641	0.91326	0.88094	0.0043019	0.0078562	0.0068941	0.012606
		Specificity	0.98202	0.92486	0.91982	0.88653	0.003811	0.0093033	0.0061701	0.012587
		Sensitivity	0.041622	0.05055	0.27409	0.38724	0.02709	0.026597	0.063725	0.087866
✗	OI 10	Accuracy	0.97355	0.94078	0.92042	0.90936	0.005592	0.0061094	0.0087236	0.0061979
		Specificity	0.98251	0.94939	0.92585	0.91437	0.004708	0.0065525	0.0074743	0.0049799
		Sensitivity	0.03373	0.045094	0.43198	0.48377	0.026694	0.027059	0.082132	0.079974

Table C.5: Post-processed results for KNN, patient 21.

Method: KNN, Neighbors: 2, 5, 10										
			Original			Veto 1			Veto 2	
KDE	Red		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
		TP	6	8	7	5	7	7	4	7	7
		FP	92	115	81	49	78	63	31	53	48
		FN	2	0	1	3	1	1	4	1	1
		TP	4	4	3	4	4	3	4	4	3
✗		FP	43	59	37	30	37	25	18	27	13
		FN	4	4	5	4	4	5	4	4	5
		TP	6	8	6	6	7	6	6	7	6
	FS 2	FP	96	129	94	67	72	72	37	51	46
		FN	2	0	2	2	1	2	2	1	2
		TP	6	8	7	6	8	7	6	8	7
	OI 2	FP	110	114	85	71	69	66	50	55	52
		FN	2	0	1	2	0	1	2	0	1
		TP	2	3	3	2	3	3	2	3	3
✗	FS 2	FP	45	57	38	30	39	28	18	23	19
		FN	6	5	5	6	5	5	6	5	5
		TP	3	3	3	3	3	3	2	3	3
✗	OI 2	FP	35	53	46	17	33	31	9	27	21
		FN	5	5	5	5	5	5	6	5	5
		TP	6	7	7	6	7	6	6	7	6
	FS 5	FP	98	128	90	57	77	65	38	53	46
		FN	2	1	1	2	1	2	2	1	2
		TP	6	8	7	6	7	6	6	7	6
	OI 5	FP	93	113	75	57	75	59	39	55	48
		FN	2	0	1	2	1	2	2	1	2
		TP	5	4	3	4	4	3	4	4	3
✗	FS 5	FP	47	56	29	28	37	21	14	21	11
		FN	3	4	5	4	4	5	4	4	5
		TP	3	4	3	3	4	3	3	4	3
✗	OI 5	FP	39	54	37	29	39	24	16	25	15
		FN	5	4	5	5	4	5	5	4	5
		TP	7	8	8	6	7	8	6	6	6
	FS 10	FP	84	120	83	57	78	62	32	56	44
		FN	1	0	0	2	1	0	2	2	2
		TP	6	7	7	5	6	7	5	6	7
	OI 10	FP	95	114	83	50	76	64	32	54	54
		FN	2	1	1	3	2	1	3	2	1
		TP	4	4	3	4	4	3	3	4	3
✗	FS 10	FP	40	57	44	22	40	32	15	30	19
		FN	4	4	5	4	4	5	5	4	5
		TP	2	4	2	2	3	2	2	3	2
✗	OI 10	FP	42	56	40	26	39	24	14	27	15

		FN	6	4	6	6	5	6	6	5	6

Table C.6: Post-processed results for random forest, patient 21.

Method: Random Forest, Trees: 10, 30, 50											
			Original			Veto 1			Veto 2		
KDE	Red		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
		TP	7	8	7	7	8	7	6	8	7
		FP	119	101	101	75	74	70	53	55	53
		FN	1	0	1	1	0	1	2	0	1
		TP	4	4	2	4	4	2	4	4	2
✗		FP	57	41	44	38	26	27	30	17	20
		FN	4	4	6	4	4	6	4	4	6
		TP	7	7	7	6	7	7	6	7	7
	FS 2	FP	124	111	109	79	72	71	56	56	49
		FN	1	1	1	2	1	1	2	1	1
		TP	8	8	8	8	8	8	8	8	8
	OI 2	FP	134	123	117	80	76	81	57	53	57
		FN	0	0	0	0	0	0	0	0	0
		TP	4	6	3	4	4	3	4	4	3
✗	FS 2	FP	60	52	44	37	31	27	28	22	20
		FN	4	2	5	4	4	5	4	4	5
		TP	4	4	4	4	4	4	3	3	3
✗	OI 2	FP	59	53	52	35	32	30	25	22	22
		FN	4	4	4	4	4	4	5	5	5
		TP	7	8	7	6	8	6	6	8	6
	FS 5	FP	121	103	103	78	75	78	58	50	59
		FN	1	0	1	2	0	2	2	0	2
		TP	7	7	7	6	6	6	6	6	6
	OI 5	FP	132	116	105	80	74	72	56	56	56
		FN	1	1	1	2	2	2	2	2	2
		TP	4	4	3	4	3	3	3	3	3
✗	FS 5	FP	56	51	45	32	31	29	22	22	21
		FN	4	4	5	4	5	5	5	5	5
		TP	5	4	4	5	2	4	5	4	4
✗	OI 5	FP	56	51	48	35	29	27	31	20	20
		FN	3	4	4	3	6	4	3	4	4
		TP	8	7	8	8	6	7	7	6	7
	FS 10	FP	118	103	101	78	77	73	58	52	53
		FN	0	1	0	0	2	1	1	2	1
		TP	8	8	8	8	7	8	8	7	8
	OI 10	FP	122	110	102	74	83	73	60	61	53
		FN	0	0	0	0	1	0	0	1	0
		TP	5	4	4	5	3	4	4	3	4
✗	FS 10	FP	57	47	43	34	30	30	24	20	21
		FN	3	4	4	3	5	4	4	5	4
		TP	3	3	4	3	2	3	3	3	3
✗	OI 10	FP	55	47	45	32	31	27	30	24	18
		FN	5	5	4	5	6	5	5	5	5

Table C.7: Post-processed results for logistic regression, patient 21.

Method: Logistic Regression					
KDE	Red		Original	Veto 1	Veto 2
		TP	8	8	7
		FP	167	85	58
		FN	0	0	1
		TP	4	3	3
✗		FP	66	42	31
		FN	4	5	5

	FS 2	TP	6	6	6
		FP	118	81	56
		FN	2	2	2
	OI 2	TP	7	7	6
		FP	104	77	54
		FN	1	1	2
✗	FS 2	TP	3	2	3
		FP	45	30	21
		FN	5	6	5
✗	OI 2	TP	3	2	3
		FP	38	23	19
		FN	5	6	5
	FS 5	TP	8	8	8
		FP	128	86	59
		FN	0	0	0
	OI 5	TP	8	7	7
		FP	107	76	55
		FN	0	1	1
✗	FS 5	TP	3	3	3
		FP	49	34	24
		FN	5	5	5
✗	OI 5	TP	3	3	3
		FP	38	24	20
		FN	5	5	5
	FS 10	TP	7	6	6
		FP	150	85	59
		FN	1	2	2
	OI 10	TP	7	7	7
		FP	117	82	58
		FN	1	1	1
✗	FS 10	TP	3	3	3
		FP	58	33	27
		FN	5	5	5
✗	OI 10	TP	4	3	3
		FP	50	33	24
		FN	4	5	5

Table C.8: Post-processed results SVM, patient 21.

Method: SVM, kernel: rbf, linear, BoxConst: 1, 100														
			Original				Veto 1				Veto 2			
KDE	Red		μ_1	μ_2	μ_3	μ_4	μ_1	μ_2	μ_3	μ_4	μ_1	μ_2	μ_3	μ_4
		TP	1	1	8	8	0	0	7	7	0	0	6	6
		FP	29	51	117	158	2	8	81	86	0	0	56	61
		FN	7	7	0	0	8	8	1	1	8	8	2	2
✗		TP	0	1	3	4	0	0	3	4	0	0	2	3
		FP	5	22	46	66	3	12	29	37	1	4	19	29
		FN	8	7	5	4	8	8	5	4	8	8	6	5
	FS 2	TP	7	6	6	6	6	4	6	6	6	4	6	6
		FP	155	165	93	97	88	80	72	73	61	44	49	48
		FN	1	2	2	2	2	4	2	2	2	4	2	2
	OI 2	TP	8	8	7	7	7	7	7	7	7	7	6	6
		FP	111	155	98	99	79	76	70	75	57	55	56	55
		FN	0	0	1	1	1	1	1	1	1	1	2	2
✗	FS 2	TP	3	3	3	3	2	3	2	2	3	2	3	3
		FP	53	60	39	36	30	43	24	23	24	26	14	16
		FN	5	5	5	5	6	5	6	6	5	6	5	5
✗	OI 2	TP	3	3	3	3	3	3	2	2	3	3	3	3
		FP	38	61	34	36	24	31	21	23	20	28	17	18
		FN	5	5	5	5	5	5	6	6	5	5	5	5
	FS 5	TP	5	7	7	7	3	5	7	6	2	4	7	5
		FP	124	150	115	123	63	75	76	74	38	39	55	56
		FN	3	1	1	1	5	3	1	2	6	4	1	3

	OI 5	TP	7	7	7	8	6	6	7	7	6	6	7	7
		FP	114	141	100	98	72	75	73	73	48	50	58	58
		FN	1	1	1	0	2	2	1	1	2	2	1	1
✗	FS 5	TP	2	3	3	4	2	2	3	3	1	2	3	3
		FP	45	64	43	45	27	42	29	25	20	30	19	18
		FN	6	5	5	4	6	6	5	5	7	6	5	5
✗	OI 5	TP	1	3	3	3	1	3	3	3	1	3	3	3
		FP	41	59	33	34	30	39	20	21	19	27	17	16
		FN	7	5	5	5	7	5	5	5	7	5	5	5
	FS 10	TP	2	2	7	8	0	1	7	7	0	0	6	6
		FP	70	87	110	148	23	28	79	88	11	9	58	61
		FN	6	6	1	0	8	7	1	1	8	8	2	2
	OI 10	TP	2	2	7	7	1	1	7	7	0	0	7	7
		FP	51	93	97	114	11	27	70	82	1	7	55	57
		FN	6	6	1	1	7	7	1	1	8	8	1	1
✗	FS 10	TP	0	2	3	3	0	1	3	3	0	1	3	3
		FP	15	44	49	56	10	29	32	34	4	19	21	20
		FN	8	6	5	5	8	7	5	5	8	7	5	5
✗	OI 10	TP	0	1	3	3	0	1	3	3	0	1	3	3
		FP	15	37	43	47	8	22	29	29	5	18	21	22
		FN	8	7	5	5	8	7	5	5	8	7	5	5

C.2 Cluster balanced training set

Table C.9: Cross-validation results for KNN, patient 21.

Method: KNN, Neighbors: 2, 5, 10								
KDE	Red		μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
		Accuracy	0.94948	0.79338	0.89592	0.0054177	0.022136	0.0138
		Specificity	0.96136	0.80138	0.90658	0.006185	0.023478	0.014836
		Sensitivity	0.051786	0.22671	0.12565	0.028048	0.05223	0.053028
✗		Accuracy	0.95067	0.78813	0.85243	0.0028668	0.018921	0.017536
		Specificity	0.95939	0.79446	0.85968	0.0039932	0.020601	0.019079
		Sensitivity	0.054122	0.18331	0.15522	0.019417	0.050229	0.043617
	FS 2	Accuracy	0.82962	0.66812	0.6934	0.027172	0.052573	0.10661
		Specificity	0.83699	0.67198	0.69679	0.027318	0.054253	0.10821
		Sensitivity	0.27866	0.436	0.48752	0.053176	0.07296	0.064549
	OI 2	Accuracy	0.89357	0.7886	0.87874	0.012368	0.023431	0.0114
		Specificity	0.90147	0.79316	0.88532	0.013316	0.02447	0.01228
		Sensitivity	0.2871	0.45976	0.39562	0.079232	0.074448	0.083435
✗	FS 2	Accuracy	0.86813	0.72746	0.81191	0.013045	0.03509	0.042836
		Specificity	0.87479	0.73175	0.8165	0.014311	0.037014	0.043901
		Sensitivity	0.24438	0.38243	0.44292	0.057257	0.080049	0.071104
✗	OI 2	Accuracy	0.90034	0.76717	0.85691	0.012928	0.021902	0.016876
		Specificity	0.90737	0.77151	0.86251	0.014431	0.023769	0.018529
		Sensitivity	0.22399	0.43359	0.34116	0.052828	0.069291	0.058321
	FS 5	Accuracy	0.88646	0.79604	0.8826	0.0074597	0.02268	0.016371
		Specificity	0.89567	0.80269	0.89066	0.0081278	0.023423	0.016754
		Sensitivity	0.2097	0.29632	0.28957	0.052721	0.07971	0.074803
	OI 5	Accuracy	0.91054	0.79105	0.89774	0.0069476	0.020342	0.010675
		Specificity	0.92023	0.79695	0.90634	0.007763	0.021634	0.01095
		Sensitivity	0.17785	0.36717	0.24278	0.051756	0.063934	0.071678
✗	FS 5	Accuracy	0.89218	0.77951	0.84365	0.015721	0.023643	0.020339
		Specificity	0.89951	0.78421	0.84945	0.016814	0.025414	0.021737
		Sensitivity	0.1847	0.33739	0.27246	0.060763	0.091237	0.069596
✗	OI 5	Accuracy	0.91794	0.78715	0.85712	0.005707	0.017037	0.014943
		Specificity	0.92561	0.79182	0.86357	0.006841	0.018949	0.016534
		Sensitivity	0.1525	0.3886	0.23915	0.051618	0.076211	0.072127

	FS 10	Accuracy	0.93102	0.80044	0.90161	0.0075825	0.020867	0.0090198
		Specificity	0.94184	0.80728	0.91108	0.008345	0.021898	0.010214
		Sensitivity	0.10909	0.29791	0.22532	0.04311	0.056287	0.064841
	OI 10	Accuracy	0.9409	0.80623	0.92236	0.0060081	0.019679	0.0095101
		Specificity	0.95182	0.81386	0.93239	0.0067813	0.020942	0.010367
		Sensitivity	0.11623	0.23957	0.17175	0.039993	0.067075	0.065591
✗	FS 10	Accuracy	0.9333	0.78945	0.85649	0.0069249	0.018807	0.019953
		Specificity	0.94119	0.79494	0.86269	0.0078744	0.020551	0.021549
		Sensitivity	0.14	0.27671	0.29532	0.040516	0.061767	0.070469
✗	OI 10	Accuracy	0.94363	0.77449	0.87177	0.0044136	0.020804	0.012367
		Specificity	0.95189	0.78006	0.87889	0.0056668	0.022568	0.014016
		Sensitivity	0.10834	0.2534	0.15701	0.041458	0.055734	0.059736

Table C.10: Cross-validation results for randf, patient 21.

Method: Random Forest, Trees: 10, 30, 50								
KDE	Red		μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
		Accuracy	0.43134	0.43154	0.43214	0.053774	0.042451	0.045728
		Specificity	0.42939	0.43002	0.43028	0.055919	0.044364	0.047436
		Sensitivity	0.59798	0.57477	0.58829	0.099378	0.097569	0.085962
✗		Accuracy	0.73718	0.75783	0.75809	0.033088	0.024512	0.027033
		Specificity	0.74011	0.76124	0.76179	0.034014	0.025732	0.027884
		Sensitivity	0.62931	0.60668	0.59351	0.055241	0.061374	0.088216
	FS 2	Accuracy	0.44619	0.63354	0.61973	0.038902	0.038013	0.037743
		Specificity	0.44513	0.635	0.62171	0.039732	0.038941	0.038481
		Sensitivity	0.50146	0.51849	0.51203	0.05387	0.080039	0.067271
	OI 2	Accuracy	0.73823	0.72447	0.76692	0.034506	0.049574	0.037622
		Specificity	0.74128	0.7276	0.77058	0.03615	0.051021	0.039481
		Sensitivity	0.48918	0.46407	0.4669	0.10477	0.084598	0.10727
✗	FS 2	Accuracy	0.7666	0.78868	0.80353	0.031363	0.030618	0.025286
		Specificity	0.77113	0.79321	0.80741	0.032437	0.031847	0.026594
		Sensitivity	0.5128	0.54471	0.62913	0.05268	0.066071	0.071112
✗	OI 2	Accuracy	0.85137	0.87279	0.87468	0.011797	0.012328	0.012923
		Specificity	0.85773	0.8786	0.8812	0.012733	0.012991	0.013326
		Sensitivity	0.41521	0.48609	0.441	0.08928	0.10682	0.089378
	FS 5	Accuracy	0.46734	0.58882	0.56266	0.05227	0.035485	0.035727
		Specificity	0.46612	0.58942	0.56302	0.053972	0.036484	0.036799
		Sensitivity	0.55352	0.54398	0.52583	0.11468	0.098189	0.11933
	OI 5	Accuracy	0.66749	0.67903	0.67457	0.03835	0.034074	0.032796
		Specificity	0.66887	0.68089	0.67686	0.039448	0.034935	0.034395
		Sensitivity	0.57466	0.54527	0.51674	0.084639	0.086769	0.096086
✗	FS 5	Accuracy	0.76561	0.78656	0.789	0.023901	0.021604	0.021978
		Specificity	0.76892	0.79016	0.79287	0.024852	0.021853	0.022544
		Sensitivity	0.59128	0.62505	0.60056	0.10094	0.068974	0.053153
✗	OI 5	Accuracy	0.8021	0.81905	0.81932	0.024798	0.019409	0.02162
		Specificity	0.80674	0.82391	0.82406	0.025324	0.019728	0.022005
		Sensitivity	0.52764	0.53024	0.54045	0.068205	0.084765	0.087243
	FS 10	Accuracy	0.44384	0.67637	0.65248	0.035222	0.022817	0.026694
		Specificity	0.44157	0.67788	0.65439	0.036216	0.024313	0.028203
		Sensitivity	0.63968	0.56061	0.51917	0.066842	0.12541	0.092641
	OI 10	Accuracy	0.4712	0.52767	0.53231	0.043285	0.051615	0.045033
		Specificity	0.46934	0.52755	0.53288	0.044249	0.053597	0.046883
		Sensitivity	0.62293	0.55581	0.52009	0.096248	0.11079	0.11413
✗	FS 10	Accuracy	0.75781	0.80187	0.8079	0.018888	0.028157	0.021173
		Specificity	0.76062	0.80578	0.81246	0.020139	0.029008	0.021706
		Sensitivity	0.64434	0.59647	0.5616	0.0775	0.090092	0.069135
✗	OI 10	Accuracy	0.79111	0.78763	0.79156	0.024093	0.022072	0.025371
		Specificity	0.79446	0.79112	0.79483	0.024706	0.022835	0.026637
		Sensitivity	0.64416	0.61967	0.66531	0.053513	0.085791	0.06455

Table C.11: Cross-validation results for logistic regression, patient 21.

Method: Logistic Regression				
KDE	Red		μ_1	σ_1
		Accuracy	0.7367	0.016892
		Specificity	0.74152	0.018098
		Sensitivity	0.41349	0.066515
✗		Accuracy	0.75869	0.016811
		Specificity	0.7629	0.018141
		Sensitivity	0.37589	0.045206
	FS 2	Accuracy	0.78564	0.01249
		Specificity	0.79109	0.012327
		Sensitivity	0.38746	0.065487
	OI 2	Accuracy	0.91765	0.013389
		Specificity	0.9255	0.013593
		Sensitivity	0.3211	0.085497
✗	FS 2	Accuracy	0.85005	0.028954
		Specificity	0.85531	0.029048
		Sensitivity	0.31925	0.078493
✗	OI 2	Accuracy	0.92829	0.015034
		Specificity	0.93612	0.015994
		Sensitivity	0.14448	0.072908
	FS 5	Accuracy	0.83409	0.020606
		Specificity	0.84048	0.020849
		Sensitivity	0.36508	0.037541
	OI 5	Accuracy	0.84852	0.0096096
		Specificity	0.8547	0.0098174
		Sensitivity	0.38395	0.060772
✗	FS 5	Accuracy	0.85806	0.017904
		Specificity	0.86423	0.018142
		Sensitivity	0.24522	0.038641
✗	OI 5	Accuracy	0.88484	0.012657
		Specificity	0.89062	0.011653
		Sensitivity	0.34178	0.064759
	FS 10	Accuracy	0.80476	0.016008
		Specificity	0.81028	0.016548
		Sensitivity	0.43171	0.087521
	OI 10	Accuracy	0.82777	0.0088898
		Specificity	0.83396	0.0089632
		Sensitivity	0.38483	0.053942
✗	FS 10	Accuracy	0.8025	0.016667
		Specificity	0.80749	0.01786
		Sensitivity	0.37574	0.080224
✗	OI 10	Accuracy	0.82879	0.015926
		Specificity	0.83438	0.01596
		Sensitivity	0.30224	0.051583

Table C.12: Cross-validation results for SVM, patient 21.

Method: SVM, Trees: 10, 30, 50										
KDE	Red		μ_1	μ_2	μ_3	μ_4	σ_1	σ_2	σ_3	σ_4
		Accuracy	0.96139	0.94622	0.78862	0.73021	0.0030997	0.0054369	0.016323	0.020412
		Specificity	0.97411	0.95862	0.79425	0.73564	0.004049	0.0063899	0.016862	0.021645
		Sensitivity	0	0.0089286	0.40853	0.36726	0	0.0089286	0.085189	0.066434
✗		Accuracy	0.97125	0.95721	0.81252	0.75937	0.0026864	0.0039074	0.016912	0.019384
		Specificity	0.98059	0.96618	0.81755	0.76367	0.0038686	0.0054972	0.017983	0.020828
		Sensitivity	0.0089286	0.032693	0.35281	0.39752	0.0089286	0.019316	0.073153	0.06121
	FS 2	Accuracy	0.83925	0.63379	0.823	0.82293	0.036369	0.040855	0.053373	0.054862
		Specificity	0.8476	0.63885	0.83105	0.83074	0.036989	0.041996	0.054895	0.056592
		Sensitivity	0.20168	0.26979	0.26883	0.23782	0.055592	0.04741	0.051346	0.081548
	OI 2	Accuracy	0.84508	0.7413	0.92871	0.92548	0.022397	0.022158	0.017144	0.017844
		Specificity	0.85032	0.745	0.93794	0.93466	0.023405	0.02283	0.01771	0.018406
		Sensitivity	0.44369	0.44163	0.21916	0.21916	0.074941	0.077291	0.085916	0.085916

✗	FS 2	Accuracy	0.79796	0.71565	0.92952	0.91576	0.028513	0.011652	0.032853	0.041302
		Specificity	0.80393	0.72017	0.93695	0.92306	0.030148	0.012308	0.033746	0.042317
		Sensitivity	0.26214	0.27489	0.19968	0.19968	0.065333	0.04175	0.078142	0.078142
✗	OI 2	Accuracy	0.84895	0.76512	0.98804	0.98232	0.014964	0.022389	0.0044534	0.0058549
		Specificity	0.85438	0.7691	0.99756	0.9916	0.016472	0.023977	0.0024427	0.0047194
		Sensitivity	0.34376	0.41574	0	0.022727	0.061931	0.06613	0	0.022727
	FS 5	Accuracy	0.88444	0.78319	0.84415	0.81316	0.016991	0.029157	0.041871	0.042606
		Specificity	0.89559	0.79173	0.85186	0.81922	0.018466	0.029981	0.04279	0.043882
		Sensitivity	0.051948	0.13122	0.28782	0.38101	0.020534	0.039932	0.081787	0.11431
	OI 5	Accuracy	0.8224	0.79647	0.87095	0.85373	0.01877	0.016734	0.012734	0.012326
		Specificity	0.83087	0.80412	0.87817	0.85931	0.019849	0.017625	0.013538	0.012893
		Sensitivity	0.1858	0.21503	0.3375	0.42859	0.056904	0.052969	0.072446	0.067938
✗	FS 5	Accuracy	0.87111	0.80295	0.82992	0.77388	0.014003	0.01294	0.02669	0.029617
		Specificity	0.87897	0.80926	0.83529	0.77883	0.015212	0.014255	0.028326	0.03222
		Sensitivity	0.074414	0.16769	0.35422	0.39286	0.03807	0.049347	0.084779	0.10199
✗	OI 5	Accuracy	0.85199	0.81905	0.91141	0.88552	0.018066	0.012098	0.023109	0.021358
		Specificity	0.85874	0.82505	0.91845	0.89164	0.019229	0.012967	0.023278	0.021587
		Sensitivity	0.1799	0.23984	0.22143	0.31088	0.049644	0.072081	0.070975	0.068704
	FS 10	Accuracy	0.94552	0.88238	0.792	0.7698	0.0062911	0.015277	0.030984	0.027778
		Specificity	0.95805	0.89361	0.798	0.77515	0.0073118	0.016506	0.031797	0.028611
		Sensitivity	0	0.038149	0.39113	0.40898	0	0.01988	0.077898	0.058468
	OI 10	Accuracy	0.9539	0.9237	0.86932	0.84728	0.0043066	0.0078991	0.011977	0.014255
		Specificity	0.96639	0.93553	0.87621	0.85401	0.0052794	0.0088626	0.012174	0.014461
		Sensitivity	0.0089286	0.029221	0.36311	0.36782	0.0089286	0.014416	0.067594	0.054176
✗	FS 10	Accuracy	0.95149	0.87617	0.8152	0.8057	0.011242	0.013247	0.018629	0.017858
		Specificity	0.96055	0.88424	0.81991	0.81019	0.012423	0.015069	0.019387	0.018821
		Sensitivity	0.023764	0.054022	0.38983	0.43145	0.013107	0.026398	0.087707	0.076897
✗	OI 10	Accuracy	0.96076	0.93847	0.87111	0.82796	0.0034273	0.0041281	0.025883	0.026501
		Specificity	0.96983	0.9471	0.87748	0.83358	0.0045706	0.0058769	0.02649	0.026773
		Sensitivity	0.023764	0.05055	0.26916	0.29321	0.013107	0.026597	0.056386	0.043479

Table C.13: Post-processed results for KNN, patient 21.

Method: KNN, Neighbors: 2, 5, 10										
			Original			Veto 1			Veto 2	
KDE	Red		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
		TP	2	6	2	0	3	2	0	3	0
		FP	45	155	92	9	45	47	4	28	26
		FN	6	2	6	8	5	6	8	5	8
✗		TP	1	5	4	0	4	3	0	4	2
		FP	34	68	61	15	33	34	2	33	25
		FN	7	3	4	8	4	5	8	4	6
	FS 2	TP	6	6	8	5	5	7	5	5	6
		FP	149	194	151	50	50	53	36	44	50
		FN	2	2	0	3	3	1	3	3	2
	OI 2	TP	7	7	6	6	5	6	6	5	5
		FP	97	148	95	48	49	54	38	43	41
		FN	1	1	2	2	3	2	2	3	3
✗	FS 2	TP	5	7	3	4	7	3	3	4	2
		FP	62	69	63	28	30	30	20	32	26
		FN	3	1	5	4	1	5	5	4	6
✗	OI 2	TP	3	5	5	3	4	5	2	3	4
		FP	54	71	65	29	32	37	17	34	30
		FN	5	3	3	5	4	3	6	5	4
	FS 5	TP	5	5	5	3	4	5	3	4	3
		FP	97	143	102	45	49	48	22	36	33
		FN	3	3	3	5	4	3	5	4	5
	OI 5	TP	5	6	6	5	4	6	1	5	3
		FP	81	153	89	36	60	50	23	36	25
		FN	3	2	2	3	4	2	7	3	5
✗	FS 5	TP	4	6	5	3	6	5	2	4	3
		FP	56	69	60	26	35	32	15	35	27
		FN	4	2	3	5	2	3	6	4	5

✗	OI 5	TP	4	6	5	3	6	5	2	6	4
		FP	41	69	61	21	34	34	10	38	25
		FN	4	2	3	5	2	3	6	2	4
	FS 10	TP	4	6	5	2	5	3	2	4	3
		FP	65	145	92	29	45	46	14	27	24
		FN	4	2	3	6	3	5	6	4	5
	OI 10	TP	4	4	4	3	4	3	1	3	2
		FP	51	145	65	14	48	41	8	30	23
		FN	4	4	4	5	4	5	7	5	6
✗	FS 10	TP	2	6	4	2	6	2	0	4	2
		FP	39	69	62	18	34	34	10	36	27
		FN	6	2	4	6	2	6	8	4	6
✗	OI 10	TP	1	5	3	0	4	2	0	4	1
		FP	39	69	58	22	34	31	7	38	26
		FN	7	3	5	8	4	6	8	4	7

Table C.14: Post-processed results for random forest, patient 21.

Method: Random Forest, Trees: 10, 30, 50											
			Original			Veto 1			Veto 2		
KDE	Red		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
		TP	6	7	7	1	4	4	5	5	5
		FP	224	240	228	35	40	40	48	50	51
		FN	2	1	1	7	4	4	3	3	3
✗		TP	6	6	5	6	6	5	6	5	4
		FP	76	75	74	31	28	29	40	37	35
		FN	2	2	3	2	2	3	2	3	4
	FS 2	TP	7	7	7	5	5	5	6	4	4
		FP	237	192	193	55	45	40	46	52	47
		FN	1	1	1	3	3	3	2	4	4
	OI 2	TP	7	7	7	6	7	7	7	5	6
		FP	157	162	144	63	59	61	49	49	46
		FN	1	1	1	2	1	1	1	3	2
✗	FS 2	TP	6	6	6	6	6	6	6	4	5
		FP	75	71	67	31	29	32	35	36	39
		FN	2	2	2	2	2	2	2	4	3
✗	OI 2	TP	6	5	5	6	5	5	4	5	5
		FP	67	68	65	39	39	39	39	37	37
		FN	2	3	3	2	3	3	4	3	3
	FS 5	TP	7	7	7	4	5	4	4	4	3
		FP	229	204	214	54	40	44	48	48	47
		FN	1	1	1	4	3	4	4	4	5
	OI 5	TP	7	7	7	5	7	3	5	5	4
		FP	177	171	178	49	47	44	52	42	47
		FN	1	1	1	3	1	5	3	3	4
✗	FS 5	TP	6	6	6	6	5	6	5	5	5
		FP	75	70	68	31	28	26	37	38	36
		FN	2	2	2	2	3	2	3	3	3
✗	OI 5	TP	6	5	6	6	5	5	5	5	5
		FP	69	67	65	32	30	30	41	38	40
		FN	2	3	2	2	3	3	3	3	3
	FS 10	TP	7	7	7	2	6	6	5	4	3
		FP	227	176	185	35	44	49	48	45	48
		FN	1	1	1	6	2	2	3	4	5
	OI 10	TP	7	6	6	4	4	3	5	5	5
		FP	230	216	207	52	44	41	53	47	52
		FN	1	2	2	4	4	5	3	3	3
✗	FS 10	TP	6	6	6	6	6	6	5	5	6
		FP	73	67	68	33	29	30	39	34	37
		FN	2	2	2	2	2	2	3	3	2
✗	OI 10	TP	6	6	6	6	6	6	5	4	5
		FP	71	71	71	36	32	37	40	39	37
		FN	2	2	2	2	2	2	3	4	3

Table C.15: Post-processed results for logistic regression, patient 21.

Method: Logistic Regression					
KDE	Red		Original	Veto 1	Veto 2
		TP	7	7	5
		FP	172	70	50
		FN	1	1	3
✗		TP	6	5	4
		FP	70	31	33
		FN	2	3	4
	FS 2	TP	6	6	5
		FP	120	68	50
		FN	2	2	3
	OI 2	TP	6	6	5
		FP	61	50	36
		FN	2	2	3
✗	FS 2	TP	4	4	3
		FP	47	22	18
		FN	4	4	5
✗	OI 2	TP	1	0	0
		FP	36	25	18
		FN	7	8	8
	FS 5	TP	7	7	5
		FP	108	66	48
		FN	1	1	3
	OI 5	TP	7	7	5
		FP	103	65	50
		FN	1	1	3
✗	FS 5	TP	4	4	2
		FP	57	30	24
		FN	4	4	6
✗	OI 5	TP	3	3	2
		FP	52	28	18
		FN	5	5	6
	FS 10	TP	8	6	6
		FP	128	67	43
		FN	0	2	2
	OI 10	TP	7	6	5
		FP	115	68	49
		FN	1	2	3
✗	FS 10	TP	5	4	4
		FP	70	29	30
		FN	3	4	4
✗	OI 10	TP	4	3	3
		FP	61	36	24
		FN	4	5	5

Table C.16: Post-processed results SVM, patient 21.

Method: SVM, kernel: rbf, linear, BoxConst: 1, 100															
			Original				Veto 1				Veto 2				
KDE	Red		μ_1	μ_2	μ_3	μ_4	μ_1	μ_2	μ_3	μ_4	μ_1	μ_2	μ_3	μ_4	
		TP	1	1	8	7	0	0	7	7	0	0	7	5	
		FP	38	55	141	174	1	5	59	64	0	0	53	52	
		FN	7	7	0	1	8	8	1	1	8	8	1	3	
✗		TP	1	1	5	5	0	0	5	5	0	0	4	5	
		FP	20	29	65	71	6	11	33	30	1	4	30	36	
		FN	7	7	3	3	8	8	3	3	8	8	4	3	
	FS 2	TP	6	6	6	5	4	3	5	4	2	3	3	2	
		FP	132	205	100	101	50	38	39	34	29	37	34	31	
		FN	2	2	2	3	4	5	3	4	6	5	5	6	

	OI 2	TP	7	7	4	4	7	7	4	4	5	5	4	4
		FP	114	174	45	49	59	60	35	38	45	42	29	32
		FN	1	1	4	4	1	1	4	4	3	3	4	4
✗	FS 2	TP	6	6	2	2	6	6	2	2	3	2	1	1
		FP	69	72	23	25	38	32	9	11	28	35	11	11
		FN	2	2	6	6	2	2	6	6	5	6	7	7
✗	OI 2	TP	4	5	0	0	4	5	0	0	4	4	0	0
		FP	63	69	4	5	30	36	4	4	28	31	1	1
		FN	4	3	8	8	4	3	8	8	4	4	8	8
	FS 5	TP	2	6	5	4	1	2	5	4	0	2	3	3
		FP	100	166	97	111	24	40	46	54	11	22	41	43
		FN	6	2	3	4	7	6	3	4	8	6	5	5
	OI 5	TP	5	5	6	7	5	4	6	7	2	3	4	5
		FP	124	136	86	93	51	57	58	61	36	42	41	43
		FN	3	3	2	1	3	4	2	1	6	5	4	3
✗	FS 5	TP	4	5	5	6	2	4	3	4	0	3	2	3
		FP	53	68	60	62	27	30	28	28	21	30	25	27
		FN	4	3	3	2	6	4	5	4	8	5	6	5
✗	OI 5	TP	4	4	3	4	4	4	3	2	2	2	2	2
		FP	63	66	37	47	31	33	20	28	22	24	14	21
		FN	4	4	5	4	4	4	5	6	6	6	6	6
	FS 10	TP	1	2	8	8	0	1	7	8	0	0	6	6
		FP	52	113	130	146	7	12	59	67	2	3	52	50
		FN	7	6	0	0	8	7	1	0	8	8	2	2
	OI 10	TP	1	1	6	6	0	0	5	5	0	0	5	5
		FP	45	75	88	102	4	14	60	59	1	1	50	49
		FN	7	7	2	2	8	8	3	3	8	8	3	3
✗	FS 10	TP	1	3	3	5	0	2	3	4	0	0	3	2
		FP	25	54	66	69	9	26	32	33	3	19	33	31
		FN	7	5	5	3	8	6	5	4	8	8	5	6
✗	OI 10	TP	1	1	4	4	0	0	3	3	0	0	2	2
		FP	29	39	53	63	12	17	28	32	4	12	18	24
		FN	7	7	4	4	8	8	5	5	8	8	6	6

D

Appendix 4: HMS Patient 36 result

D.1 Randomly balanced training set

Table D.1: Cross-validation results for KNN, patient 36.

Method: KNN, Neighbors: 2, 5, 10								
KDE	Red		μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
		Accuracy	0.89091	0.80844	0.8601	0.018529	0.025594	0.021151
		Specificity	0.89677	0.80988	0.86268	0.020553	0.02716	0.021679
		Sensitivity	0.43805	0.6398	0.61184	0.17166	0.18854	0.20066
		Accuracy	0.9438	0.88493	0.9168	0.0058793	0.0095656	0.013432
✗		Specificity	0.94797	0.88699	0.91942	0.0055488	0.0093758	0.012979
		Sensitivity	0.39309	0.5773	0.5318	0.10182	0.15939	0.18695
		Accuracy	0.84767	0.77627	0.82093	0.032677	0.045048	0.046386
	FS 2	Specificity	0.85179	0.7777	0.82426	0.033929	0.044772	0.045967
		Sensitivity	0.52193	0.65132	0.57127	0.1092	0.097197	0.075588
		Accuracy	0.85035	0.74624	0.76974	0.027996	0.016411	0.033789
	OI 2	Specificity	0.85792	0.75056	0.77355	0.02822	0.016521	0.032562
		Sensitivity	0.31689	0.41393	0.45175	0.03236	0.1021	0.18768
		Accuracy	0.92883	0.88886	0.89783	0.009588	0.017463	0.018515
✗	FS 2	Specificity	0.93315	0.89103	0.90001	0.0099142	0.017634	0.018637
		Sensitivity	0.36184	0.60252	0.58443	0.093026	0.039156	0.13553
		Accuracy	0.92174	0.85542	0.89359	0.0057592	0.016406	0.0028406
✗	OI 2	Specificity	0.92659	0.85826	0.89756	0.0056097	0.016699	0.0025176
		Sensitivity	0.28509	0.44846	0.35088	0.091705	0.15504	0.15864
		Accuracy	0.8684	0.77768	0.82973	0.027276	0.040189	0.029142
	FS 5	Specificity	0.87425	0.779	0.83377	0.029235	0.040834	0.028812
		Sensitivity	0.41721	0.63322	0.51151	0.16328	0.13987	0.11647
		Accuracy	0.86693	0.7873	0.80433	0.03964	0.045607	0.050473
	OI 5	Specificity	0.87376	0.79152	0.8067	0.03986	0.046712	0.050717
		Sensitivity	0.36897	0.43805	0.58443	0.053669	0.14037	0.15248
		Accuracy	0.93773	0.88248	0.90403	0.004977	0.014156	0.013127
✗	FS 5	Specificity	0.94205	0.8847	0.90588	0.0049989	0.014706	0.012495
		Sensitivity	0.36513	0.57072	0.61897	0.10362	0.11075	0.17813
		Accuracy	0.92365	0.86737	0.88497	0.0077742	0.011416	0.0065257
✗	OI 5	Specificity	0.92815	0.86966	0.8881	0.0077079	0.011681	0.0058225
		Sensitivity	0.32675	0.52138	0.43421	0.11221	0.18197	0.19417
		Accuracy	0.8776	0.78409	0.82221	0.012993	0.038627	0.016149
	FS 10	Specificity	0.88254	0.78567	0.82586	0.014587	0.039625	0.014976
		Sensitivity	0.47971	0.61568	0.52522	0.17286	0.15237	0.14959
		Accuracy	0.89378	0.82162	0.86381	0.022105	0.034707	0.021738
	OI 10	Specificity	0.89948	0.8227	0.86716	0.023295	0.035597	0.022544
		Sensitivity	0.46272	0.68531	0.59156	0.11352	0.1625	0.11193
		Accuracy	0.93754	0.87307	0.91251	0.0034955	0.014857	0.014719
✗	FS 10	Specificity	0.94108	0.87476	0.91452	0.0035322	0.014531	0.014357
		Sensitivity	0.45559	0.62281	0.60855	0.12286	0.13221	0.17364
		Accuracy	0.94812	0.8927	0.92262	0.0060072	0.0067555	0.010029
✗	OI 10	Specificity	0.9525	0.89453	0.92538	0.0062763	0.0069988	0.0098632

		Sensitivity	0.35088	0.60855	0.52138	0.15021	0.17364	0.18726

Table D.2: Cross-validation results for randf, patient 36.

Method: Random Forest, Trees: 10, 30, 50								
KDE	Red		μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
		Accuracy	0.80508	0.83286	0.84106	0.03649	0.038385	0.033255
		Specificity	0.80669	0.83469	0.84338	0.03796	0.039865	0.03481
		Sensitivity	0.63322	0.64364	0.60855	0.14669	0.14787	0.18098
✗		Accuracy	0.90357	0.93615	0.92737	0.016575	0.013834	0.018054
		Specificity	0.90565	0.93889	0.92965	0.017104	0.014384	0.018549
		Sensitivity	0.59485	0.54605	0.58059	0.13982	0.14598	0.18525
	FS 2	Accuracy	0.75894	0.77159	0.79999	0.039635	0.015987	0.026303
		Specificity	0.76171	0.77379	0.80278	0.040049	0.016359	0.027083
		Sensitivity	0.54331	0.60252	0.57072	0.044095	0.039156	0.10155
	OI 2	Accuracy	0.79176	0.81186	0.82423	0.041543	0.046405	0.052379
		Specificity	0.7946	0.81522	0.82827	0.041847	0.046364	0.052379
		Sensitivity	0.54989	0.5466	0.51535	0.12176	0.089041	0.071782
✗	FS 2	Accuracy	0.86858	0.91249	0.91097	0.011388	0.0077652	0.0079663
		Specificity	0.87138	0.91536	0.914	0.01235	0.0077873	0.0087056
		Sensitivity	0.51206	0.52193	0.49397	0.092229	0.10307	0.1411
✗	OI 2	Accuracy	0.88307	0.88972	0.88453	0.028925	0.025822	0.030098
		Specificity	0.8863	0.89246	0.88721	0.029023	0.025649	0.029878
		Sensitivity	0.4205	0.50439	0.50439	0.1773	0.14056	0.15062
	FS 5	Accuracy	0.81421	0.83444	0.83498	0.040035	0.015704	0.012393
		Specificity	0.81578	0.83779	0.83873	0.041344	0.01678	0.013396
		Sensitivity	0.64364	0.54605	0.52522	0.14787	0.15145	0.13591
	OI 5	Accuracy	0.81548	0.8283	0.82537	0.053309	0.053576	0.054851
		Specificity	0.81804	0.832	0.82915	0.054381	0.054113	0.055739
		Sensitivity	0.56689	0.52193	0.49013	0.17664	0.11642	0.18112
✗	FS 5	Accuracy	0.90421	0.90525	0.91309	0.018364	0.0097422	0.0076859
		Specificity	0.90704	0.90793	0.91587	0.019181	0.010688	0.0077339
		Sensitivity	0.50439	0.51809	0.53235	0.15062	0.1739	0.10865
✗	OI 5	Accuracy	0.86582	0.8892	0.89476	0.0021458	0.014717	0.0019674
		Specificity	0.86859	0.89141	0.89718	0.0023541	0.014672	0.0025
		Sensitivity	0.48355	0.56689	0.55318	0.11166	0.15293	0.12698
	FS 10	Accuracy	0.79489	0.81729	0.83035	0.034196	0.035044	0.044676
		Specificity	0.79747	0.82059	0.83359	0.034973	0.035948	0.045837
		Sensitivity	0.55647	0.52851	0.54605	0.15344	0.16001	0.14259
	OI 10	Accuracy	0.80493	0.84479	0.84514	0.045621	0.031171	0.03713
		Specificity	0.80597	0.84718	0.84767	0.047029	0.032558	0.038906
		Sensitivity	0.6716	0.61568	0.59101	0.14928	0.15658	0.19363
✗	FS 10	Accuracy	0.90082	0.91079	0.92002	0.017385	0.013754	0.011469
		Specificity	0.90297	0.9135	0.92244	0.01849	0.014592	0.011922
		Sensitivity	0.58443	0.51096	0.55976	0.1618	0.18019	0.17485
✗	OI 10	Accuracy	0.91212	0.91984	0.9267	0.015872	0.019483	0.019395
		Specificity	0.91398	0.92235	0.92878	0.016842	0.020256	0.019841
		Sensitivity	0.61897	0.54934	0.62281	0.18878	0.17784	0.12847

Table D.3: Cross-validation results for logistic regression, patient 36.

Method: Logistic Regression				
KDE	Red		μ_1	σ_1
		Accuracy	0.8051	0.0081387
		Specificity	0.80947	0.0093905
		Sensitivity	0.44134	0.1702
✗		Accuracy	0.86293	0.017029
		Specificity	0.8648	0.017973
		Sensitivity	0.59156	0.13798
	FS 2	Accuracy	0.87571	0.0058954
		Specificity	0.88062	0.0074659
		Sensitivity	0.5011	0.12932

	OI 2	Accuracy	0.80788	0.0079679
		Specificity	0.81271	0.0069156
		Sensitivity	0.45614	0.084451
✗	FS 2	Accuracy	0.9029	0.026967
		Specificity	0.90636	0.028104
		Sensitivity	0.43476	0.14972
✗	OI 2	Accuracy	0.91934	0.010348
		Specificity	0.92363	0.011086
		Sensitivity	0.36897	0.11245
	FS 5	Accuracy	0.82266	0.016883
		Specificity	0.82635	0.017469
		Sensitivity	0.50768	0.17382
	OI 5	Accuracy	0.79452	0.0392
		Specificity	0.79722	0.041084
		Sensitivity	0.5318	0.18695
✗	FS 5	Accuracy	0.90515	0.020637
		Specificity	0.90816	0.021794
		Sensitivity	0.46217	0.22396
✗	OI 5	Accuracy	0.90168	0.017137
		Specificity	0.90417	0.017692
		Sensitivity	0.56743	0.085973
	FS 10	Accuracy	0.79855	0.034549
		Specificity	0.80396	0.034424
		Sensitivity	0.38268	0.11233
	OI 10	Accuracy	0.81323	0.040394
		Specificity	0.81526	0.042329
		Sensitivity	0.60855	0.17643
✗	FS 10	Accuracy	0.88847	0.019054
		Specificity	0.89165	0.019843
		Sensitivity	0.43805	0.19476
✗	OI 10	Accuracy	0.91705	0.012847
		Specificity	0.91983	0.013774
		Sensitivity	0.51096	0.19072

Table D.4: Cross-validation results for SVM, patient 36.

Method: SVM, Trees: 10, 30, 50										
KDE	Red		μ_1	μ_2	μ_3	μ_4	σ_1	σ_2	σ_3	σ_4
		Accuracy	0.98434	0.98294	0.86451	0.7887	0.0032723	0.004038	0.0062893	0.016338
		Specificity	0.99821	0.99679	0.86903	0.79195	0.0011063	0.0018865	0.0078564	0.018052
		Sensitivity	0.020833	0.020833	0.483	0.50055	0.010417	0.010417	0.19418	0.17716
✗		Accuracy	0.99137	0.9904	0.91186	0.87998	0.001697	0.0024086	0.017421	0.02604
		Specificity	0.99912	0.99814	0.91476	0.88241	0.00061419	0.0014485	0.018471	0.027339
		Sensitivity	0.020833	0.020833	0.483	0.51096	0.010417	0.010417	0.19915	0.19825
	FS 2	Accuracy	0.71195	0.73222	0.88469	0.85349	0.030148	0.029104	0.0095234	0.020358
		Specificity	0.71021	0.73374	0.88934	0.85786	0.030604	0.028895	0.010218	0.023397
		Sensitivity	0.8114	0.61678	0.52193	0.51864	0.063596	0.099508	0.12582	0.17891
	OI 2	Accuracy	0.78953	0.75593	0.77015	0.76035	0.044851	0.015649	0.035977	0.031547
		Specificity	0.79385	0.76013	0.77438	0.76387	0.045541	0.01574	0.035579	0.030452
		Sensitivity	0.44518	0.43147	0.46656	0.49781	0.1187	0.0884	0.094866	0.12611
✗	FS 2	Accuracy	0.8366	0.85063	0.89058	0.88576	0.015293	0.0089402	0.039695	0.030681
		Specificity	0.8366	0.85095	0.89352	0.88844	0.015935	0.0088332	0.041351	0.031887
		Sensitivity	0.8114	0.7977	0.47643	0.52906	0.07312	0.054409	0.1897	0.15736
✗	OI 2	Accuracy	0.86046	0.8583	0.90192	0.90174	0.019284	0.027904	0.027594	0.02721
		Specificity	0.86362	0.86142	0.90659	0.90631	0.019511	0.028402	0.028278	0.027785
		Sensitivity	0.42434	0.42434	0.2955	0.30592	0.11283	0.11983	0.151	0.1476
	FS 5	Accuracy	0.61828	0.64878	0.89016	0.82659	0.033575	0.050482	0.0085353	0.01511
		Specificity	0.61403	0.6462	0.89545	0.83121	0.034962	0.052735	0.0093625	0.016777
		Sensitivity	0.88103	0.80811	0.47643	0.45559	0.072657	0.080488	0.14039	0.16961
	OI 5	Accuracy	0.83089	0.81996	0.77461	0.77863	0.076676	0.060259	0.053833	0.048427
		Specificity	0.83691	0.82565	0.77628	0.78074	0.078451	0.061284	0.056071	0.05044
		Sensitivity	0.358	0.38268	0.58059	0.55263	0.12924	0.087945	0.18525	0.19737
		Accuracy	0.86734	0.84795	0.9081	0.90014	0.060635	0.043692	0.028064	0.031954

✗	FS 5	Specificity	0.8703	0.84969	0.91177	0.90282	0.063297	0.045578	0.028515	0.033778
		Sensitivity	0.41667	0.53125	0.4068	0.48629	0.30046	0.27776	0.12575	0.25559
✗	OI 5	Accuracy	0.91075	0.89845	0.88923	0.88639	0.010967	0.014422	0.028592	0.017453
		Specificity	0.91462	0.90232	0.89193	0.88943	0.01156	0.014941	0.028967	0.017653
		Sensitivity	0.36458	0.35417	0.49342	0.45175	0.18251	0.17708	0.21402	0.17696
	FS 10	Accuracy	0.6922	0.77516	0.88113	0.80885	0.13742	0.12443	0.0046639	0.013758
		Specificity	0.6909	0.77715	0.88571	0.8136	0.14266	0.12936	0.0065307	0.013641
		Sensitivity	0.625	0.51042	0.49342	0.4205	0.31302	0.28086	0.20706	0.1773
	OI 10	Accuracy	0.81643	0.81565	0.84797	0.7998	0.14576	0.13802	0.020412	0.044633
		Specificity	0.82221	0.82125	0.85133	0.80245	0.1518	0.14383	0.022403	0.046912
		Sensitivity	0.34375	0.35417	0.53509	0.54934	0.28356	0.2797	0.21792	0.17784
✗	FS 10	Accuracy	0.97252	0.96446	0.914	0.89399	0.0091078	0.010713	0.025807	0.024181
		Specificity	0.97933	0.97101	0.91681	0.89682	0.0090405	0.010649	0.027141	0.025311
		Sensitivity	0.10417	0.125	0.49342	0.47259	0.05512	0.065052	0.21402	0.20462
✗	OI 10	Accuracy	0.96691	0.96147	0.91638	0.91331	0.017259	0.016617	0.028096	0.016956
		Specificity	0.97375	0.96788	0.91893	0.91559	0.017186	0.016467	0.029288	0.017864
		Sensitivity	0.09375	0.13542	0.53893	0.5773	0.065052	0.072917	0.18318	0.17681

Table D.5: Post-processed results for KNN, patient 36.

Method: KNN, Neighbors: 2, 5, 10										
			Original			Veto 1			Veto 2	
KDE	Red		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
		TP	3	3	3	3	3	3	3	3	3
		FP	82	124	100	56	78	61	34	47	36
		FN	0	0	0	0	0	0	0	0	0
✗		TP	2	2	3	1	2	2	1	2	2
		FP	14	25	24	9	19	13	6	13	8
		FN	1	1	0	2	1	1	2	1	1
	FS 2	TP	3	3	3	3	3	3	3	3	3
		FP	117	131	112	76	85	80	40	62	54
		FN	0	0	0	0	0	0	0	0	0
	OI 2	TP	3	3	3	3	3	3	3	3	3
		FP	111	162	154	61	81	74	34	54	47
		FN	0	0	0	0	0	0	0	0	0
✗	FS 2	TP	2	3	3	2	3	3	1	2	2
		FP	20	21	23	14	15	16	11	12	9
		FN	1	0	0	1	0	0	2	1	1
✗	OI 2	TP	3	3	3	2	3	3	0	3	3
		FP	19	24	22	13	17	16	10	12	12
		FN	0	0	0	1	0	0	3	0	0
	FS 5	TP	3	3	3	3	3	3	3	3	3
		FP	96	135	119	63	80	67	41	55	44
		FN	0	0	0	0	0	0	0	0	0
	OI 5	TP	3	3	3	3	3	3	3	3	3
		FP	95	128	119	63	80	74	40	56	53
		FN	0	0	0	0	0	0	0	0	0
✗	FS 5	TP	2	3	3	2	1	2	1	1	1
		FP	17	27	22	11	19	15	6	11	9
		FN	1	0	0	1	2	1	2	2	2
✗	OI 5	TP	3	3	2	3	3	2	2	2	2
		FP	19	22	26	15	17	19	10	14	14
		FN	0	0	1	0	0	1	1	1	1
	FS 10	TP	3	3	3	3	3	3	3	3	3
		FP	95	133	132	60	76	65	37	52	40
		FN	0	0	0	0	0	0	0	0	0
	OI 10	TP	3	3	3	3	3	3	3	3	3
		FP	83	115	97	56	79	65	31	54	44
		FN	0	0	0	0	0	0	0	0	0
✗	FS 10	TP	2	2	2	1	2	2	1	1	0
		FP	14	27	22	11	20	16	6	15	10
		FN	1	1	1	2	1	1	2	2	3
✗	OI 10	TP	3	3	3	3	3	3	1	2	1
		FP	15	25	20	9	17	13	3	14	8

		FN	0	0	0	0	0	0	2	1	2

Table D.6: Post-processed results for random forest, patient 36.

Method: Random Forest, Trees: 10, 30, 50											
			Original			Veto 1			Veto 2		
KDE	Red		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
		TP	3	3	3	3	3	3	3	3	3
		FP	111	101	93	80	74	71	57	56	52
		FN	0	0	0	0	0	0	0	0	0
		TP	2	2	1	2	1	1	1	1	1
✗		FP	21	17	17	14	10	9	7	4	4
		FN	1	1	2	1	2	2	2	2	2
		TP	3	3	3	3	3	3	3	3	3
	FS 2	FP	135	126	125	84	83	76	54	68	54
		FN	0	0	0	0	0	0	0	0	0
		TP	3	3	3	3	3	3	3	3	3
	OI 2	FP	123	108	101	75	73	70	50	52	49
		FN	0	0	0	0	0	0	0	0	0
		TP	3	2	2	2	2	2	2	1	2
✗	FS 2	FP	22	19	20	15	13	14	12	11	10
		FN	0	1	1	1	1	1	1	2	1
		TP	2	3	2	2	2	2	2	2	2
✗	OI 2	FP	21	23	20	16	15	16	10	9	10
		FN	1	0	1	1	1	1	1	1	1
		TP	3	3	3	3	3	3	3	3	3
	FS 5	FP	112	115	111	82	78	69	58	49	49
		FN	0	0	0	0	0	0	0	0	0
		TP	3	3	3	3	3	3	3	3	3
	OI 5	FP	104	102	103	73	69	70	50	48	51
		FN	0	0	0	0	0	0	0	0	0
		TP	3	2	2	1	2	2	1	2	2
✗	FS 5	FP	21	21	19	15	15	12	9	9	11
		FN	0	1	1	2	1	1	2	1	1
		TP	3	3	3	3	2	2	1	2	2
✗	OI 5	FP	25	24	24	16	14	16	13	9	11
		FN	0	0	0	0	1	1	2	1	1
		TP	3	3	3	3	3	3	3	3	3
	FS 10	FP	124	115	107	80	76	73	57	48	47
		FN	0	0	0	0	0	0	0	0	0
		TP	3	3	3	3	3	3	3	3	3
	OI 10	FP	100	97	93	80	73	69	56	54	53
		FN	0	0	0	0	0	0	0	0	0
		TP	2	2	2	1	2	2	1	2	2
✗	FS 10	FP	22	22	18	16	14	11	11	9	6
		FN	1	1	1	2	1	1	2	1	1
		TP	3	2	2	2	1	1	1	1	1
✗	OI 10	FP	19	18	17	14	10	12	10	8	8
		FN	0	1	1	1	2	2	2	2	2

Table D.7: Post-processed results for logistic regression, patient 36.

Method: Logistic Regression					
KDE	Red		Original	Veto 1	Veto 2
		TP	3	3	3
		FP	127	75	44
		FN	0	0	0
		TP	3	2	2
✗		FP	25	16	9
		FN	0	1	1

	FS 2	TP	3	3	3
		FP	89	54	35
		FN	0	0	0
	OI 2	TP	3	3	3
		FP	117	68	50
		FN	0	0	0
✗	FS 2	TP	2	2	2
		FP	19	11	8
		FN	1	1	1
✗	OI 2	TP	1	1	1
		FP	23	13	8
		FN	2	2	2
	FS 5	TP	3	3	3
		FP	116	74	46
		FN	0	0	0
	OI 5	TP	3	3	3
		FP	111	78	58
		FN	0	0	0
✗	FS 5	TP	2	2	2
		FP	19	14	6
		FN	1	1	1
✗	OI 5	TP	3	3	3
		FP	23	10	5
		FN	0	0	0
	FS 10	TP	3	3	3
		FP	126	73	37
		FN	0	0	0
	OI 10	TP	3	3	3
		FP	116	76	53
		FN	0	0	0
✗	FS 10	TP	3	3	2
		FP	20	14	9
		FN	0	0	1
✗	OI 10	TP	3	2	1
		FP	18	15	11
		FN	0	1	2

Table D.8: Post-processed results SVM, patient 36.

Method: SVM, kernel: rbf, linear, BoxConst: 1, 100														
			Original				Veto 1				Veto 2			
KDE	Red		μ_1	μ_2	μ_3	μ_4	μ_1	μ_2	μ_3	μ_4	μ_1	μ_2	μ_3	μ_4
		TP	0	0	3	3	0	0	3	3	0	0	3	3
		FP	0	0	94	127	0	0	61	78	0	0	36	44
		FN	3	3	0	0	3	3	0	0	3	3	0	0
✗		TP	0	0	3	3	0	0	2	3	0	0	2	2
		FP	0	2	20	23	0	1	14	17	0	1	10	12
		FN	3	3	0	0	3	3	1	0	3	3	1	1
	FS 2	TP	3	3	3	3	3	3	3	3	3	3	3	3
		FP	163	149	81	94	85	87	57	57	64	65	42	38
		FN	0	0	0	0	0	0	0	0	0	0	0	0
	OI 2	TP	3	3	3	3	3	3	3	3	3	3	3	3
		FP	130	165	130	138	75	81	74	75	55	55	59	60
		FN	0	0	0	0	0	0	0	0	0	0	0	0
✗	FS 2	TP	3	3	2	2	3	3	2	2	3	2	2	2
		FP	27	27	17	21	17	16	11	12	11	13	10	8
		FN	0	0	1	1	0	0	1	1	0	1	1	1
✗	OI 2	TP	3	3	2	2	2	3	2	2	2	1	2	2
		FP	22	22	24	24	17	17	13	14	12	14	6	7
		FN	0	0	1	1	1	0	1	1	1	2	1	1
	FS 5	TP	3	3	3	3	3	3	3	3	3	3	3	3
		FP	195	193	70	115	88	92	53	68	68	61	31	47
		FN	0	0	0	0	0	0	0	0	0	0	0	0

	OI 5	TP	3	3	3	3	3	3	3	3	3	3	3	3
		FP	102	118	113	111	67	72	81	78	43	49	63	65
		FN	0	0	0	0	0	0	0	0	0	0	0	0
✗	FS 5	TP	3	3	2	2	2	3	2	2	1	1	1	2
		FP	20	22	19	17	14	16	13	13	11	14	9	9
		FN	0	0	1	1	1	0	1	1	2	2	2	1
✗	OI 5	TP	3	3	3	3	2	2	3	3	2	1	3	3
		FP	18	18	24	28	15	16	14	11	9	12	7	5
		FN	0	0	0	0	1	1	0	0	1	2	0	0
	FS 10	TP	2	2	3	3	2	2	3	3	2	2	3	3
		FP	134	117	87	122	69	57	55	77	51	34	35	40
		FN	1	1	0	0	1	1	0	0	1	1	0	0
	OI 10	TP	2	2	3	3	1	1	3	3	1	2	3	3
		FP	73	83	86	117	51	59	64	81	37	37	46	50
		FN	1	1	0	0	2	2	0	0	2	1	0	0
✗	FS 10	TP	0	0	3	3	0	0	2	2	0	0	1	1
		FP	6	10	15	16	4	8	13	14	3	5	6	7
		FN	3	3	0	0	3	3	1	1	3	3	2	2
✗	OI 10	TP	0	0	1	2	0	0	1	1	0	0	1	1
		FP	8	10	17	18	6	7	10	15	5	5	8	10
		FN	3	3	2	1	3	3	2	2	3	3	2	2

D.2 Cluster balanced training set

Table D.9: Cross-validation results for KNN, patient 36.

Method: KNN, Neighbors: 2, 5, 10								
KDE	Red		μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
		Accuracy	0.96766	0.86379	0.89308	0.0016176	0.00030949	0.011448
		Specificity	0.9792	0.87085	0.90097	0.0027707	0.0016609	0.010887
		Sensitivity	0.14583	0.33004	0.30208	0.075116	0.15636	0.17148
✗		Accuracy	0.9603	0.84874	0.8296	0.0097117	0.046857	0.0659
		Specificity	0.96665	0.85183	0.83337	0.010195	0.046614	0.066193
		Sensitivity	0.14583	0.41009	0.30208	0.075116	0.17711	0.15763
	FS 2	Accuracy	0.82243	0.67251	0.73477	0.041137	0.030701	0.035667
		Specificity	0.82975	0.67522	0.73816	0.04133	0.032974	0.037734
		Sensitivity	0.26042	0.40625	0.42708	0.1378	0.21271	0.22268
	OI 2	Accuracy	0.90467	0.75982	0.73557	0.016817	0.017492	0.02978
		Specificity	0.91351	0.76574	0.74179	0.016988	0.017717	0.030304
		Sensitivity	0.25713	0.31634	0.26042	0.10264	0.10592	0.13052
✗	FS 2	Accuracy	0.91859	0.81828	0.7652	0.011285	0.026287	0.048967
		Specificity	0.92327	0.82089	0.76792	0.01133	0.026913	0.049481
		Sensitivity	0.3273	0.43805	0.37171	0.070826	0.14037	0.15954
✗	OI 2	Accuracy	0.90509	0.78333	0.73796	0.0064694	0.040237	0.06274
		Specificity	0.91039	0.78618	0.74107	0.0061495	0.040608	0.063347
		Sensitivity	0.22259	0.37884	0.30208	0.058662	0.13708	0.15131
	FS 5	Accuracy	0.9151	0.78683	0.8198	0.014312	0.005557	0.025218
		Specificity	0.92354	0.79138	0.82667	0.013719	0.0053423	0.026697
		Sensitivity	0.28838	0.41338	0.30208	0.12971	0.18832	0.18072
	OI 5	Accuracy	0.91746	0.82962	0.83629	0.029181	0.015645	0.023055
		Specificity	0.92521	0.83593	0.84346	0.029565	0.01638	0.024175
		Sensitivity	0.33717	0.34046	0.29167	0.1163	0.1442	0.15023
✗	FS 5	Accuracy	0.92983	0.82078	0.82944	0.015349	0.033739	0.025736
		Specificity	0.93428	0.82322	0.83304	0.014987	0.034185	0.025846
		Sensitivity	0.33004	0.45175	0.33004	0.14889	0.1815	0.14332
✗	OI 5	Accuracy	0.93078	0.80604	0.76039	0.0073098	0.018291	0.021678
		Specificity	0.93564	0.80854	0.76279	0.0072672	0.017739	0.021644
		Sensitivity	0.28838	0.44134	0.40296	0.12327	0.17951	0.18088

	FS 10	Accuracy	0.94799	0.84196	0.85907	0.00077955	0.0073314	0.031902
		Specificity	0.95747	0.84712	0.8655	0.00097504	0.0073348	0.031365
		Sensitivity	0.25	0.43092	0.37171	0.1263	0.16383	0.16847
	OI 10	Accuracy	0.96032	0.86706	0.90312	0.0023594	0.0012416	0.010252
		Specificity	0.97	0.87344	0.91078	0.0032289	0.0026535	0.010205
		Sensitivity	0.25713	0.36458	0.32292	0.11175	0.19292	0.16764
✗	FS 10	Accuracy	0.94176	0.84725	0.8137	0.0097024	0.014252	0.041819
		Specificity	0.94633	0.85015	0.81697	0.009258	0.014175	0.041919
		Sensitivity	0.33717	0.43092	0.35088	0.13203	0.17159	0.15864
✗	OI 10	Accuracy	0.9594	0.86642	0.86122	0.0087782	0.037094	0.041203
		Specificity	0.96516	0.86956	0.86548	0.0092571	0.037029	0.041343
		Sensitivity	0.20833	0.40625	0.28125	0.11024	0.21271	0.14321

Table D.10: Cross-validation results for randf, patient 36.

Method: Random Forest, Trees: 10, 30, 50								
KDE	Red		μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
		Accuracy	0.68425	0.74592	0.75466	0.079087	0.066359	0.078802
		Specificity	0.68397	0.74757	0.75601	0.081185	0.068909	0.0819
		Sensitivity	0.63322	0.54221	0.55592	0.13514	0.19237	0.22605
✗		Accuracy	0.93839	0.93604	0.93857	0.0084714	0.01918	0.01843
		Specificity	0.94419	0.94249	0.94522	0.010292	0.020896	0.01977
		Sensitivity	0.46546	0.42379	0.41338	0.21586	0.19603	0.18395
	FS 2	Accuracy	0.80955	0.76381	0.73662	0.035185	0.050771	0.042495
		Specificity	0.81283	0.76797	0.74112	0.036761	0.051719	0.043495
		Sensitivity	0.51809	0.43147	0.37555	0.15615	0.095481	0.12582
	OI 2	Accuracy	0.7444	0.77705	0.80167	0.015457	0.039325	0.040627
		Specificity	0.74846	0.78159	0.80726	0.015438	0.040295	0.041466
		Sensitivity	0.42434	0.4068	0.358	0.11983	0.12575	0.12669
✗	FS 2	Accuracy	0.9133	0.93205	0.93165	0.014891	0.015086	0.022921
		Specificity	0.91948	0.93955	0.93685	0.015327	0.016323	0.024625
		Sensitivity	0.45943	0.37555	0.52522	0.074052	0.13578	0.14959
✗	OI 2	Accuracy	0.93786	0.94193	0.93674	0.013728	0.0097462	0.014465
		Specificity	0.94556	0.95041	0.94478	0.01335	0.009654	0.014736
		Sensitivity	0.37226	0.31634	0.33717	0.085275	0.10592	0.1163
	FS 5	Accuracy	0.66759	0.76279	0.71602	0.13072	0.059134	0.039381
		Specificity	0.66962	0.76827	0.72144	0.13392	0.060096	0.040096
		Sensitivity	0.42379	0.31963	0.28125	0.1858	0.13621	0.14768
	OI 5	Accuracy	0.78533	0.81063	0.79394	0.043115	0.046585	0.063122
		Specificity	0.7909	0.81563	0.79896	0.043034	0.04768	0.064122
		Sensitivity	0.36897	0.4068	0.38596	0.069524	0.12575	0.11545
✗	FS 5	Accuracy	0.92095	0.93265	0.92878	0.019189	0.024155	0.026817
		Specificity	0.92688	0.93958	0.93564	0.020607	0.026115	0.028461
		Sensitivity	0.47314	0.39254	0.39967	0.14446	0.19241	0.16811
✗	OI 5	Accuracy	0.93728	0.94138	0.94833	0.016462	0.0099939	0.0057867
		Specificity	0.94328	0.94935	0.9564	0.017718	0.01024	0.0055454
		Sensitivity	0.48355	0.35471	0.35471	0.13173	0.10206	0.098823
	FS 10	Accuracy	0.76866	0.77555	0.79346	0.077716	0.077804	0.068804
		Specificity	0.77105	0.77977	0.7983	0.079715	0.079784	0.0698
		Sensitivity	0.54276	0.41721	0.40351	0.11909	0.14422	0.11064
	OI 10	Accuracy	0.75985	0.77253	0.78951	0.062551	0.060335	0.065275
		Specificity	0.76213	0.7748	0.79294	0.064303	0.062778	0.066972
		Sensitivity	0.53564	0.5318	0.49397	0.15016	0.1904	0.12391
✗	FS 10	Accuracy	0.92584	0.94119	0.93941	0.019733	0.017455	0.019916
		Specificity	0.93094	0.94805	0.94608	0.021582	0.018413	0.020493
		Sensitivity	0.51809	0.40296	0.42763	0.1739	0.1754	0.13487
✗	OI 10	Accuracy	0.93266	0.94211	0.93715	0.018078	0.012006	0.022536
		Specificity	0.93858	0.94829	0.94311	0.020315	0.013827	0.024599
		Sensitivity	0.46217	0.4375	0.46217	0.20419	0.22535	0.19606

Table D.11: Cross-validation results for logistic regression, patient 36.

Method: Logistic Regression				
KDE	Red		μ_1	σ_1
		Accuracy	0.8088	0.0048214
		Specificity	0.81415	0.0050596
		Sensitivity	0.38925	0.14314
✗		Accuracy	0.88317	0.013083
		Specificity	0.88687	0.012672
		Sensitivity	0.38596	0.12623
	FS 2	Accuracy	0.72912	0.088712
		Specificity	0.73559	0.090401
		Sensitivity	0.26425	0.081517
	OI 2	Accuracy	0.77189	0.058319
		Specificity	0.7796	0.059575
		Sensitivity	0.22588	0.11275
✗	FS 2	Accuracy	0.75493	0.10807
		Specificity	0.75637	0.10962
		Sensitivity	0.483	0.18909
✗	OI 2	Accuracy	0.81621	0.046538
		Specificity	0.82005	0.04797
		Sensitivity	0.29879	0.17001
	FS 5	Accuracy	0.83114	0.023855
		Specificity	0.83696	0.024695
		Sensitivity	0.37884	0.13944
	OI 5	Accuracy	0.81554	0.028704
		Specificity	0.81985	0.030406
		Sensitivity	0.45175	0.17696
✗	FS 5	Accuracy	0.80705	0.017621
		Specificity	0.80907	0.01685
		Sensitivity	0.5148	0.16024
✗	OI 5	Accuracy	0.81426	0.016926
		Specificity	0.81753	0.016676
		Sensitivity	0.36129	0.15459
	FS 10	Accuracy	0.79436	0.01609
		Specificity	0.7987	0.017551
		Sensitivity	0.43092	0.17159
	OI 10	Accuracy	0.79262	0.015873
		Specificity	0.79596	0.01778
		Sensitivity	0.50055	0.18259
✗	FS 10	Accuracy	0.87874	0.026567
		Specificity	0.88243	0.026713
		Sensitivity	0.36842	0.14228
✗	OI 10	Accuracy	0.86376	0.012789
		Specificity	0.86662	0.012651
		Sensitivity	0.47314	0.10842

Table D.12: Cross-validation results for SVM, patient 36.

Method: SVM, Trees: 10, 30, 50										
KDE	Red		μ_1	μ_2	μ_3	μ_4	σ_1	σ_2	σ_3	σ_4
		Accuracy	0.98521	0.9838	0.85436	0.83657	0.0029009	0.0037977	0.010806	0.025182
		Specificity	0.99927	0.99766	0.85941	0.84307	0.00072582	0.0018479	0.011893	0.026161
		Sensitivity	0.010417	0.020833	0.44134	0.32292	0.010417	0.010417	0.17399	0.16171
✗		Accuracy	0.99166	0.99108	0.8914	0.87703	0.0015343	0.0018608	0.018094	0.016204
		Specificity	0.99961	0.99883	0.89463	0.88131	0.00039273	0.00078054	0.018411	0.016277
		Sensitivity	0	0.020833	0.42379	0.29167	0	0.010417	0.19098	0.16271
	FS 2	Accuracy	0.80663	0.81041	0.76585	0.76201	0.035364	0.042093	0.086069	0.079595
		Specificity	0.81006	0.81761	0.77309	0.76843	0.035849	0.043937	0.088549	0.08141
		Sensitivity	0.47917	0.26042	0.25384	0.2955	0.25022	0.18779	0.11005	0.10527
	OI 2	Accuracy	0.85383	0.7854	0.71816	0.71226	0.023954	0.028666	0.073773	0.075939
		Specificity	0.86141	0.79035	0.7253	0.71947	0.024253	0.029452	0.075192	0.077735
		Sensitivity	0.2955	0.40351	0.21546	0.20504	0.095547	0.098164	0.085795	0.093472

✗	FS 2	Accuracy	0.86038	0.88194	0.82645	0.81915	0.011111	0.012872	0.07122	0.075881
		Specificity	0.86474	0.88715	0.82985	0.82344	0.010846	0.012234	0.072285	0.077105
		Sensitivity	0.27796	0.19792	0.35417	0.25	0.12136	0.128	0.19874	0.14434
✗	OI 2	Accuracy	0.85889	0.82812	0.80036	0.81116	0.019848	0.026843	0.083277	0.071454
		Specificity	0.86233	0.83168	0.80477	0.81567	0.020196	0.027012	0.085229	0.073264
		Sensitivity	0.37884	0.3443	0.21875	0.21875	0.13708	0.099943	0.17399	0.17399
	FS 5	Accuracy	0.95771	0.89166	0.77782	0.80189	0.0046049	0.01841	0.061243	0.011805
		Specificity	0.96997	0.90256	0.78332	0.8079	0.0041963	0.018197	0.06157	0.011379
		Sensitivity	0.09375	0.11458	0.36842	0.34046	0.047735	0.057998	0.13158	0.1442
	OI 5	Accuracy	0.88982	0.86265	0.81489	0.81803	0.05478	0.057222	0.060881	0.052389
		Specificity	0.89678	0.86864	0.81941	0.82259	0.056455	0.058961	0.063794	0.054908
		Sensitivity	0.35088	0.38213	0.43421	0.43421	0.15343	0.16696	0.21104	0.20398
✗	FS 5	Accuracy	0.95226	0.91454	0.83501	0.8554	0.0033059	0.0042577	0.072215	0.037453
		Specificity	0.95892	0.92022	0.83794	0.85928	0.0025811	0.0044722	0.072916	0.037566
		Sensitivity	0.10417	0.17708	0.42434	0.34101	0.052083	0.089	0.10691	0.065241
✗	OI 5	Accuracy	0.93575	0.92388	0.79967	0.81123	0.0082626	0.0019986	0.024222	0.033967
		Specificity	0.94035	0.92808	0.80274	0.81479	0.0083209	0.0018134	0.023697	0.034059
		Sensitivity	0.3125	0.34375	0.36458	0.33004	0.15729	0.17211	0.18429	0.13987
	FS 10	Accuracy	0.97955	0.96603	0.83327	0.79012	0.0035072	0.0049508	0.021483	0.03997
		Specificity	0.99299	0.97875	0.83764	0.79588	0.0017413	0.0043606	0.022862	0.041197
		Sensitivity	0.041667	0.072917	0.46217	0.33004	0.02756	0.037558	0.17868	0.14332
	OI 10	Accuracy	0.978	0.96814	0.81553	0.80744	0.004024	0.0053338	0.027013	0.022631
		Specificity	0.9916	0.98088	0.81898	0.81112	0.0018756	0.0037346	0.028784	0.024257
		Sensitivity	0.03125	0.072917	0.50384	0.49013	0.018042	0.037558	0.20112	0.18112
✗	FS 10	Accuracy	0.9843	0.97644	0.87636	0.86741	0.0021902	0.0048549	0.030661	0.018474
		Specificity	0.99171	0.9834	0.87964	0.87142	0.002416	0.0049559	0.030973	0.018468
		Sensitivity	0.052083	0.09375	0.41009	0.31963	0.02756	0.047735	0.15663	0.14091
✗	OI 10	Accuracy	0.98793	0.9841	0.88407	0.85253	0.0016959	0.0024161	0.036097	0.02354
		Specificity	0.99555	0.9914	0.88688	0.85539	0.0010646	0.0023255	0.036162	0.023435
		Sensitivity	0.03125	0.0625	0.47259	0.44846	0.018042	0.036084	0.18717	0.1497

Table D.13: Post-processed results for KNN, patient 36.

Method: KNN, Neighbors: 2, 5, 10										
			Original			Veto 1			Veto 2	
KDE	Red		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
		TP	3	3	3	3	3	2	2	3	2
		FP	14	101	83	8	47	39	2	15	18
		FN	0	0	0	0	0	1	1	0	1
✗		TP	1	3	3	1	3	2	0	2	1
		FP	12	24	29	7	15	17	5	13	14
		FN	2	0	0	2	0	1	3	1	2
	FS 2	TP	3	3	3	3	3	3	3	3	3
		FP	114	165	148	45	58	53	25	41	32
		FN	0	0	0	0	0	0	0	0	0
	OI 2	TP	3	3	3	2	2	2	2	2	2
		FP	65	152	157	33	58	52	12	33	24
		FN	0	0	0	1	1	1	1	1	1
✗	FS 2	TP	3	3	3	3	3	3	0	0	2
		FP	25	27	29	12	18	17	7	14	14
		FN	0	0	0	0	0	0	3	3	1
✗	OI 2	TP	2	3	3	1	3	3	0	2	2
		FP	24	28	29	16	17	18	10	15	17
		FN	1	0	0	2	0	0	3	1	1
	FS 5	TP	3	3	3	3	3	3	3	2	3
		FP	67	135	119	37	61	44	20	33	22
		FN	0	0	0	0	0	0	0	1	0
	OI 5	TP	3	3	3	2	2	2	2	3	2
		FP	65	117	115	38	58	49	13	27	25
		FN	0	0	0	1	1	1	1	0	1
✗	FS 5	TP	2	3	3	2	3	3	1	2	1
		FP	15	29	28	8	20	17	7	15	15
		FN	1	0	0	1	0	0	2	1	2

✗	OI 5	TP	3	3	3	1	3	3	0	3	3
		FP	17	29	30	12	18	18	11	15	16
		FN	0	0	0	2	0	0	3	0	0
	FS 10	TP	3	3	3	2	2	3	3	3	3
		FP	42	112	96	21	50	31	6	25	12
		FN	0	0	0	1	1	0	0	0	0
	OI 10	TP	3	3	3	3	3	3	3	3	3
		FP	26	105	74	15	44	36	5	18	17
		FN	0	0	0	0	0	0	0	0	0
✗	FS 10	TP	2	3	3	1	3	3	0	2	3
		FP	15	28	31	10	17	17	8	14	16
		FN	1	0	0	2	0	0	3	1	0
✗	OI 10	TP	1	2	3	0	1	3	0	0	0
		FP	14	25	26	7	14	17	5	11	14
		FN	2	1	0	3	2	0	3	3	3

Table D.14: Post-processed results for random forest, patient 36.

Method: Random Forest, Trees: 10, 30, 50											
			Original			Veto 1			Veto 2		
KDE	Red		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
		TP	3	3	3	3	3	3	3	3	3
		FP	164	137	130	73	73	64	51	51	46
		FN	0	0	0	0	0	0	0	0	0
✗		TP	3	3	3	3	3	3	3	2	3
		FP	18	16	17	15	15	16	14	10	12
		FN	0	0	0	0	0	0	0	1	0
	FS 2	TP	3	3	3	3	3	3	3	2	3
		FP	118	123	146	72	65	64	43	38	41
		FN	0	0	0	0	0	0	0	1	0
	OI 2	TP	3	3	3	3	3	3	3	2	2
		FP	140	133	126	69	62	68	43	39	40
		FN	0	0	0	0	0	0	0	1	1
✗	FS 2	TP	3	3	3	3	3	3	3	2	2
		FP	20	20	19	16	15	15	16	13	11
		FN	0	0	0	0	0	0	0	1	1
✗	OI 2	TP	3	3	3	3	3	3	3	2	2
		FP	19	18	19	16	14	15	11	10	12
		FN	0	0	0	0	0	0	0	1	1
	FS 5	TP	3	3	3	3	3	2	3	3	2
		FP	158	136	148	53	62	60	43	41	39
		FN	0	0	0	0	0	1	0	0	1
	OI 5	TP	3	3	3	3	3	3	3	3	2
		FP	132	123	124	65	67	65	35	42	38
		FN	0	0	0	0	0	0	0	0	1
✗	FS 5	TP	3	3	3	3	3	3	3	2	3
		FP	19	15	15	17	13	12	15	12	12
		FN	0	0	0	0	0	0	0	1	0
✗	OI 5	TP	3	3	3	2	3	2	2	2	3
		FP	18	19	19	15	15	15	11	10	9
		FN	0	0	0	1	0	1	1	1	0
	FS 10	TP	3	3	3	3	3	3	3	3	2
		FP	124	118	107	66	61	64	42	41	40
		FN	0	0	0	0	0	0	0	0	1
	OI 10	TP	3	3	3	3	3	3	3	3	2
		FP	128	130	121	72	70	68	49	47	46
		FN	0	0	0	0	0	0	0	0	1
✗	FS 10	TP	3	3	3	3	2	3	2	2	2
		FP	19	17	17	16	14	13	15	10	11
		FN	0	0	0	0	1	0	1	1	1
✗	OI 10	TP	3	3	3	2	3	2	3	3	3
		FP	19	17	18	17	16	15	15	14	14
		FN	0	0	0	1	0	1	0	0	0

Table D.15: Post-processed results for logistic regression, patient 36.

Method: Logistic Regression					
KDE	Red		Original	Veto 1	Veto 2
		TP	3	3	3
		FP	142	69	39
		FN	0	0	0
✗		TP	3	2	2
		FP	23	16	11
		FN	0	1	1
	FS 2	TP	3	3	3
		FP	143	46	31
		FN	0	0	0
	OI 2	TP	3	2	3
		FP	116	41	28
		FN	0	1	0
✗	FS 2	TP	3	3	3
		FP	20	11	13
		FN	0	0	0
✗	OI 2	TP	2	2	1
		FP	27	16	10
		FN	1	1	2
	FS 5	TP	3	2	2
		FP	128	66	34
		FN	0	1	1
	OI 5	TP	3	2	2
		FP	118	75	41
		FN	0	1	1
✗	FS 5	TP	3	3	3
		FP	29	16	15
		FN	0	0	0
✗	OI 5	TP	3	3	3
		FP	29	16	12
		FN	0	0	0
	FS 10	TP	3	3	3
		FP	145	76	44
		FN	0	0	0
	OI 10	TP	3	3	3
		FP	139	73	44
		FN	0	0	0
✗	FS 10	TP	3	2	1
		FP	23	17	12
		FN	0	1	2
✗	OI 10	TP	3	2	1
		FP	24	16	13
		FN	0	1	2

Table D.16: Post-processed results SVM, patient 36.

Method: SVM, kernel: rbf, linear, BoxConst: 1, 100														
			Original				Veto 1				Veto 2			
KDE	Red		μ_1	μ_2	μ_3	μ_4	μ_1	μ_2	μ_3	μ_4	μ_1	μ_2	μ_3	μ_4
		TP	0	0	3	3	0	0	3	3	0	0	3	3
		FP	0	1	105	126	0	0	54	58	0	0	31	32
		FN	3	3	0	0	3	3	0	0	3	3	0	0
✗		TP	0	0	3	3	0	0	2	2	0	0	1	0
		FP	0	1	24	24	0	0	16	16	0	0	13	12
		FN	3	3	0	0	3	3	1	1	3	3	2	3
	FS 2	TP	3	3	3	3	3	3	2	2	3	2	1	2
		FP	114	124	125	125	55	61	38	37	39	33	20	28
		FN	0	0	0	0	0	0	1	1	0	1	2	1

	OI 2	TP	3	3	3	3	2	3	2	2	2	3	2	2
		FP	100	136	139	139	47	61	45	46	25	36	32	34
		FN	0	0	0	0	1	0	1	1	1	0	1	1
✗	FS 2	TP	3	3	2	2	2	0	2	2	2	0	2	2
		FP	21	23	23	23	14	15	13	12	10	12	11	12
		FN	0	0	1	1	1	3	1	1	1	3	1	1
✗	OI 2	TP	2	3	1	1	2	3	1	1	1	2	1	1
		FP	25	27	24	24	18	20	14	13	16	16	11	11
		FN	1	0	2	2	1	0	2	2	2	1	2	2
	FS 5	TP	2	3	3	3	2	2	3	2	1	2	3	3
		FP	27	75	129	126	13	24	47	60	2	8	34	36
		FN	1	0	0	0	1	1	0	1	2	1	0	0
	OI 5	TP	3	3	3	3	3	3	2	2	3	3	2	2
		FP	72	89	107	119	43	47	59	71	22	28	38	39
		FN	0	0	0	0	0	0	1	1	0	0	1	1
✗	FS 5	TP	1	2	2	3	1	2	2	2	0	0	1	2
		FP	11	23	25	26	7	13	16	17	5	11	9	15
		FN	2	1	1	0	2	1	1	1	3	3	2	1
✗	OI 5	TP	1	1	3	3	0	0	3	3	0	0	3	3
		FP	14	15	30	29	10	10	17	17	7	9	15	14
		FN	2	2	0	0	3	3	0	0	3	3	0	0
	FS 10	TP	1	2	3	3	1	2	3	2	0	0	3	3
		FP	2	19	115	144	1	4	59	58	0	0	31	39
		FN	2	1	0	0	2	1	0	1	3	3	0	0
	OI 10	TP	1	2	3	3	1	1	3	2	0	1	3	2
		FP	6	19	113	126	1	9	60	73	1	2	35	41
		FN	2	1	0	0	2	2	0	1	3	2	0	1
✗	FS 10	TP	0	0	2	3	0	0	1	2	0	0	0	0
		FP	3	7	25	27	1	3	13	18	1	3	11	14
		FN	3	3	1	0	3	3	2	1	3	3	3	3
✗	OI 10	TP	0	0	3	3	0	0	2	2	0	0	1	2
		FP	1	6	25	24	0	2	17	15	0	1	11	12
		FN	3	3	0	0	3	3	1	1	3	3	2	1

E

Appendix 5: HMS Patient 37 result

E.1 Randomly balanced training set

Table E.1: Cross-validation results for KNN, patient 37.

Method: KNN, Neighbors: 2, 5, 10								
KDE	Red		μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
		Accuracy	0.82877	0.77425	0.80689	0.015675	0.013553	0.014539
		Specificity	0.84978	0.78267	0.82173	0.02096	0.014818	0.016943
		Sensitivity	0.42687	0.66055	0.5834	0.039792	0.038355	0.0473
✗		Accuracy	0.86818	0.82039	0.85595	0.014397	0.0098436	0.0098351
		Specificity	0.8939	0.83523	0.8747	0.016244	0.010276	0.010341
		Sensitivity	0.36131	0.56087	0.5155	0.048261	0.049638	0.050264
	FS 2	Accuracy	0.81256	0.76828	0.78545	0.0171	0.014746	0.018786
		Specificity	0.83275	0.77469	0.79491	0.021409	0.017324	0.021922
		Sensitivity	0.44742	0.71251	0.70903	0.039611	0.034009	0.039448
	OI 2	Accuracy	0.82041	0.74592	0.76693	0.015135	0.018545	0.019577
		Specificity	0.84431	0.74954	0.7724	0.018063	0.020642	0.022748
		Sensitivity	0.44652	0.6982	0.70132	0.041414	0.035313	0.045204
✗	FS 2	Accuracy	0.86134	0.80914	0.83554	0.012887	0.011284	0.011985
		Specificity	0.88336	0.82198	0.85316	0.014059	0.012282	0.013254
		Sensitivity	0.44082	0.59344	0.53911	0.036685	0.035863	0.040972
✗	OI 2	Accuracy	0.85971	0.80484	0.82015	0.013594	0.015623	0.014309
		Specificity	0.88736	0.82086	0.84015	0.014467	0.0173	0.0161
		Sensitivity	0.37433	0.57493	0.5389	0.04815	0.038514	0.040769
	FS 5	Accuracy	0.82989	0.76221	0.80423	0.015002	0.015586	0.017284
		Specificity	0.84743	0.76303	0.82004	0.01889	0.019398	0.020086
		Sensitivity	0.49202	0.68389	0.60627	0.040451	0.042722	0.048489
	OI 5	Accuracy	0.82742	0.76794	0.79836	0.019215	0.01741	0.016896
		Specificity	0.84505	0.77306	0.80914	0.025314	0.018605	0.018463
		Sensitivity	0.47164	0.68754	0.63947	0.038448	0.03924	0.04067
✗	FS 5	Accuracy	0.86683	0.82251	0.83997	0.01394	0.012554	0.0114
		Specificity	0.89025	0.83464	0.85697	0.01524	0.012981	0.011889
		Sensitivity	0.41823	0.60493	0.56531	0.044764	0.045185	0.042667
✗	OI 5	Accuracy	0.87082	0.83039	0.85258	0.013038	0.01107	0.010925
		Specificity	0.89299	0.84324	0.87011	0.013985	0.010655	0.010455
		Sensitivity	0.45018	0.62993	0.57447	0.039079	0.042067	0.045114
	FS 10	Accuracy	0.83416	0.77028	0.80947	0.015135	0.013005	0.015533
		Specificity	0.85628	0.7778	0.81975	0.0186	0.013739	0.017733
		Sensitivity	0.43983	0.67077	0.63945	0.0454	0.041985	0.053553
	OI 10	Accuracy	0.82502	0.77052	0.79929	0.017633	0.015458	0.016128
		Specificity	0.84231	0.77294	0.80948	0.022409	0.017191	0.017471
		Sensitivity	0.49026	0.68862	0.6171	0.037817	0.044216	0.047302
✗	FS 10	Accuracy	0.86663	0.82449	0.84644	0.013702	0.011572	0.012815
		Specificity	0.89072	0.83522	0.86478	0.014869	0.012169	0.012025
		Sensitivity	0.40002	0.60834	0.53458	0.044183	0.047738	0.044536
✗	OI 10	Accuracy	0.86918	0.82297	0.85233	0.011681	0.011053	0.010782
		Specificity	0.89409	0.83562	0.87001	0.012337	0.011022	0.010613

		Sensitivity	0.42702	0.60884	0.5599	0.035811	0.043673	0.052301

Table E.2: Cross-validation results for randf, patient 37.

Method: Random Forest, Trees: 10, 30, 50								
KDE	Red		μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
		Accuracy	0.8069	0.82789	0.82566	0.014928	0.012558	0.013175
		Specificity	0.81053	0.83779	0.83571	0.016033	0.015263	0.014201
		Sensitivity	0.75495	0.69822	0.69325	0.045268	0.052531	0.061628
		Accuracy	0.8534	0.87125	0.8776	0.01081	0.011456	0.0096834
✗		Specificity	0.86459	0.8826	0.89017	0.011734	0.01215	0.010824
		Sensitivity	0.6733	0.67096	0.67046	0.057076	0.05717	0.057494
		Accuracy	0.75542	0.77698	0.75855	0.015055	0.015026	0.014971
	FS 2	Specificity	0.75279	0.78428	0.77142	0.016119	0.01653	0.017784
		Sensitivity	0.73541	0.67894	0.63294	0.041742	0.046472	0.038939
		Accuracy	0.75163	0.75844	0.77087	0.017755	0.016284	0.016088
	OI 2	Specificity	0.75852	0.76452	0.77523	0.022069	0.02023	0.019135
		Sensitivity	0.72509	0.71784	0.74818	0.039076	0.039707	0.038277
		Accuracy	0.82095	0.83414	0.82741	0.0086883	0.01219	0.011561
✗	FS 2	Specificity	0.82863	0.84669	0.84151	0.010198	0.015057	0.01331
		Sensitivity	0.6647	0.62794	0.60398	0.044288	0.050367	0.041664
		Accuracy	0.81405	0.81543	0.81993	0.015362	0.015436	0.014944
✗	OI 2	Specificity	0.82409	0.83059	0.83363	0.016636	0.017699	0.016616
		Sensitivity	0.64558	0.59088	0.58804	0.03921	0.044857	0.044719
		Accuracy	0.78486	0.78778	0.77942	0.015413	0.016665	0.015958
	FS 5	Specificity	0.78379	0.79619	0.78819	0.01687	0.019694	0.016882
		Sensitivity	0.78006	0.70632	0.67748	0.046927	0.039955	0.040268
		Accuracy	0.76755	0.7874	0.78811	0.015323	0.016447	0.015371
	OI 5	Specificity	0.77243	0.79305	0.7966	0.016879	0.019253	0.017597
		Sensitivity	0.71344	0.73932	0.68652	0.035889	0.038776	0.042269
		Accuracy	0.84242	0.84447	0.85144	0.010601	0.011388	0.011136
✗	FS 5	Specificity	0.85195	0.85655	0.86502	0.011078	0.013164	0.01227
		Sensitivity	0.63852	0.65753	0.60293	0.05643	0.046625	0.051274
		Accuracy	0.8347	0.84701	0.85066	0.012451	0.012571	0.01286
✗	OI 5	Specificity	0.84534	0.85868	0.86328	0.014034	0.013737	0.014364
		Sensitivity	0.67261	0.63144	0.62958	0.03874	0.040531	0.043818
		Accuracy	0.79763	0.79364	0.78682	0.013572	0.015874	0.016476
	FS 10	Specificity	0.80246	0.80138	0.79724	0.017056	0.019187	0.018632
		Sensitivity	0.74384	0.74287	0.70226	0.054549	0.041151	0.037726
		Accuracy	0.79742	0.81028	0.81524	0.016548	0.015854	0.014925
	OI 10	Specificity	0.80497	0.81606	0.82304	0.018338	0.018225	0.016792
		Sensitivity	0.6889	0.73546	0.72173	0.049854	0.048231	0.053317
		Accuracy	0.84563	0.85862	0.84934	0.01113	0.011643	0.012477
✗	FS 10	Specificity	0.85531	0.86939	0.86126	0.012064	0.012563	0.014017
		Sensitivity	0.66861	0.67271	0.65599	0.056256	0.052941	0.043722
		Accuracy	0.84396	0.86398	0.86114	0.01239	0.01202	0.011174
✗	OI 10	Specificity	0.85173	0.87535	0.87231	0.013195	0.013505	0.011991
		Sensitivity	0.70102	0.66973	0.65757	0.045327	0.045713	0.050395

Table E.3: Cross-validation results for logistic regression, patient 37.

Method: Logistic Regression				
KDE	Red		μ_1	σ_1
		Accuracy	0.80341	0.011541
		Specificity	0.81778	0.011775
		Sensitivity	0.62335	0.045828
		Accuracy	0.85098	0.010546
✗		Specificity	0.86821	0.010765
		Sensitivity	0.59178	0.046259
		Accuracy	0.77676	0.015028
	FS 2	Specificity	0.78325	0.017283

		Sensitivity	0.66882	0.056639
	OI 2	Accuracy	0.76843	0.019494
		Specificity	0.77676	0.024051
		Sensitivity	0.6646	0.065314
✗	FS 2	Accuracy	0.83143	0.01471
		Specificity	0.8477	0.016632
		Sensitivity	0.55141	0.05237
✗	OI 2	Accuracy	0.8481	0.012738
		Specificity	0.86993	0.015311
		Sensitivity	0.48009	0.065963
	FS 5	Accuracy	0.78311	0.015534
		Specificity	0.79403	0.017646
		Sensitivity	0.63624	0.05222
	OI 5	Accuracy	0.77091	0.015482
		Specificity	0.77741	0.017377
		Sensitivity	0.66265	0.052911
✗	FS 5	Accuracy	0.84024	0.013317
		Specificity	0.85493	0.01548
		Sensitivity	0.58171	0.05162
✗	OI 5	Accuracy	0.82743	0.013186
		Specificity	0.84019	0.014555
		Sensitivity	0.61377	0.053705
	FS 10	Accuracy	0.78823	0.012799
		Specificity	0.80064	0.013562
		Sensitivity	0.65555	0.038635
	OI 10	Accuracy	0.78741	0.014432
		Specificity	0.80057	0.015842
		Sensitivity	0.61086	0.053646
✗	FS 10	Accuracy	0.84684	0.011866
		Specificity	0.86297	0.012501
		Sensitivity	0.55662	0.051232
✗	OI 10	Accuracy	0.84104	0.012692
		Specificity	0.85626	0.013826
		Sensitivity	0.58476	0.056198

Table E.4: Cross-validation results for SVM, patient 37.

Method: SVM, Trees: 10, 30, 50										
KDE	Red		μ_1	μ_2	μ_3	μ_4	σ_1	σ_2	σ_3	σ_4
		Accuracy	0.56902	0.58073	0.79665	0.79397	0.017028	0.015815	0.012687	0.011788
		Specificity	0.54655	0.55788	0.81258	0.81131	0.017823	0.01638	0.013068	0.013173
		Sensitivity	0.85929	0.8608	0.61751	0.6197	0.031204	0.031107	0.047503	0.044114
✗		Accuracy	0.69728	0.70726	0.85135	0.84893	0.022158	0.02034	0.010804	0.009801
		Specificity	0.69415	0.70565	0.8699	0.86515	0.026144	0.02407	0.010893	0.0095103
		Sensitivity	0.76776	0.74603	0.57611	0.6123	0.064446	0.063315	0.050698	0.04311
	FS 2	Accuracy	0.76494	0.76943	0.75787	0.76016	0.015794	0.011639	0.017936	0.017774
		Specificity	0.77469	0.77743	0.76569	0.76754	0.020227	0.01381	0.021645	0.021583
		Sensitivity	0.64196	0.66253	0.66284	0.66955	0.060798	0.042161	0.056097	0.056869
	OI 2	Accuracy	0.73973	0.75609	0.74855	0.7485	0.018965	0.014416	0.020939	0.02096
		Specificity	0.73554	0.76196	0.75166	0.75162	0.02198	0.01712	0.025577	0.025603
		Sensitivity	0.80616	0.6913	0.7107	0.7107	0.052442	0.035892	0.063781	0.063781
✗	FS 2	Accuracy	0.81513	0.81116	0.82994	0.82691	0.013037	0.013414	0.016205	0.015986
		Specificity	0.82602	0.82171	0.84655	0.84272	0.014916	0.014154	0.018632	0.018294
		Sensitivity	0.62556	0.64148	0.55001	0.56303	0.058511	0.04097	0.061654	0.060141
✗	OI 2	Accuracy	0.81681	0.81823	0.84008	0.84117	0.015828	0.014103	0.0141	0.013781
		Specificity	0.82734	0.83433	0.85907	0.86009	0.01865	0.01571	0.016431	0.016054
		Sensitivity	0.66231	0.55503	0.49908	0.50371	0.056094	0.045795	0.068784	0.067466
	FS 5	Accuracy	0.70676	0.69115	0.77187	0.7691	0.016118	0.014732	0.015191	0.016178
		Specificity	0.70284	0.6857	0.78264	0.78148	0.018335	0.016245	0.016435	0.017544
		Sensitivity	0.75578	0.75664	0.64404	0.63617	0.040356	0.034006	0.04851	0.044102
	OI 5	Accuracy	0.76747	0.7498	0.76087	0.75747	0.01656	0.013766	0.016098	0.016559
		Specificity	0.77265	0.76062	0.76253	0.76012	0.017855	0.014542	0.018282	0.018709
		Sensitivity	0.68983	0.65894	0.71369	0.69127	0.042143	0.040053	0.054575	0.054322

✗	FS 5	Accuracy	0.75956	0.7536	0.8327	0.82932	0.015834	0.014875	0.013159	0.012979
		Specificity	0.76	0.75609	0.84652	0.8416	0.017162	0.016395	0.014269	0.013573
		Sensitivity	0.73808	0.70526	0.59858	0.60488	0.049015	0.04226	0.05429	0.056998
✗	OI 5	Accuracy	0.83256	0.8156	0.82074	0.81812	0.01194	0.012378	0.013058	0.012903
		Specificity	0.8453	0.83076	0.83362	0.83123	0.011868	0.01205	0.014544	0.014401
		Sensitivity	0.61967	0.58114	0.60087	0.6031	0.048389	0.041311	0.058213	0.05823
	FS 10	Accuracy	0.61905	0.6338	0.78342	0.79161	0.017627	0.016899	0.012391	0.012517
		Specificity	0.60393	0.62148	0.7984	0.80434	0.019576	0.019136	0.013203	0.013188
		Sensitivity	0.81364	0.79749	0.58679	0.66281	0.037563	0.038802	0.04981	0.045038
	OI 10	Accuracy	0.6521	0.66253	0.77836	0.77438	0.016529	0.016166	0.014933	0.01425
		Specificity	0.63709	0.6502	0.79285	0.78916	0.018121	0.017868	0.016319	0.015996
		Sensitivity	0.8235	0.80719	0.60223	0.59424	0.037768	0.039905	0.054946	0.055203
✗	FS 10	Accuracy	0.70495	0.71139	0.85061	0.84615	0.016397	0.015367	0.011759	0.01115
		Specificity	0.69874	0.70903	0.87013	0.86476	0.01755	0.017425	0.012974	0.012144
		Sensitivity	0.788	0.76974	0.5505	0.55859	0.049713	0.042966	0.061753	0.044892
✗	OI 10	Accuracy	0.72754	0.74251	0.83922	0.83784	0.016561	0.014851	0.012084	0.011937
		Specificity	0.72218	0.74054	0.85502	0.85338	0.017809	0.015942	0.013022	0.012874
		Sensitivity	0.79864	0.76187	0.55939	0.56532	0.043187	0.042636	0.059231	0.057123

Table E.5: Post-processed results for KNN, patient 37.

Method: KNN, Neighbors: 2, 5, 10										
			Original			Veto 1			Veto 2	
KDE	Red		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
		TP	23	24	24	20	19	20	18	18	16
		FP	76	113	90	52	60	53	27	32	30
		FN	1	0	0	4	5	4	6	6	8
✗		TP	11	16	15	9	12	12	5	10	7
		FP	53	82	66	25	34	27	12	17	13
		FN	13	8	9	15	12	12	19	14	17
	FS 2	TP	22	24	24	18	20	20	15	16	19
		FP	84	109	97	50	57	58	27	38	34
		FN	2	0	0	6	4	4	9	8	5
	OI 2	TP	24	24	23	21	19	17	15	18	18
		FP	88	114	104	45	48	55	25	39	36
		FN	0	0	1	3	5	7	9	6	6
✗	FS 2	TP	16	17	14	12	12	9	9	10	8
		FP	70	83	76	28	35	29	14	21	18
		FN	8	7	10	12	12	15	15	14	16
✗	OI 2	TP	13	14	12	9	7	8	6	5	5
		FP	72	86	79	24	28	30	16	21	22
		FN	11	10	12	15	17	16	18	19	19
	FS 5	TP	24	24	24	19	19	20	15	19	18
		FP	77	108	95	40	59	51	24	39	30
		FN	0	0	0	5	5	4	9	5	6
	OI 5	TP	22	24	24	16	19	19	15	18	16
		FP	76	104	89	39	59	50	19	33	28
		FN	2	0	0	8	5	5	9	6	8
✗	FS 5	TP	11	15	13	8	9	10	5	7	8
		FP	64	79	71	26	33	32	13	18	18
		FN	13	9	11	16	15	14	19	17	16
✗	OI 5	TP	15	16	15	9	12	11	7	8	7
		FP	55	69	62	21	29	21	11	16	13
		FN	9	8	9	15	12	13	17	16	17
	FS 10	TP	23	24	23	19	19	19	17	18	17
		FP	76	103	90	48	66	52	24	29	28
		FN	1	0	1	5	5	5	7	6	7
	OI 10	TP	24	24	23	20	20	21	18	17	17
		FP	78	109	92	48	47	50	24	31	25
		FN	0	0	1	4	4	3	6	7	7
✗	FS 10	TP	10	14	14	8	13	10	7	9	7
		FP	58	80	70	24	27	28	9	17	14
		FN	14	10	10	16	11	14	17	15	17

✗	OI 10	TP	12	15	14	9	12	11	5	8	8
		FP	58	80	67	23	32	28	11	20	13
		FN	12	9	10	15	12	13	19	16	16

Table E.6: Post-processed results for random forest, patient 37.

Method: Random Forest, Trees: 10, 30, 50											
			Original			Veto 1			Veto 2		
KDE	Red		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
		TP	24	23	22	20	19	18	17	18	18
		FP	86	82	79	46	49	45	32	25	27
		FN	0	1	2	4	5	6	7	6	6
✗		TP	15	14	16	10	11	11	8	8	9
		FP	63	59	57	24	20	23	14	17	15
		FN	9	10	8	14	13	13	16	16	15
	FS 2	TP	24	24	24	20	20	19	18	18	16
		FP	114	109	108	56	54	55	35	30	38
		FN	0	0	0	4	4	5	6	6	8
	OI 2	TP	24	24	24	16	18	18	17	18	19
		FP	110	109	105	52	50	48	39	34	36
		FN	0	0	0	8	6	6	7	6	5
✗	FS 2	TP	16	13	16	11	8	11	9	7	10
		FP	81	75	75	30	28	29	22	20	21
		FN	8	11	8	13	16	13	15	17	14
✗	OI 2	TP	14	13	13	10	9	8	8	6	7
		FP	78	76	74	24	24	25	22	20	22
		FN	10	11	11	14	15	16	16	18	17
	FS 5	TP	23	24	24	18	20	19	17	18	15
		FP	98	97	97	49	54	52	33	33	37
		FN	1	0	0	6	4	5	7	6	9
	OI 5	TP	24	24	23	18	19	20	16	18	17
		FP	102	98	99	56	51	55	41	32	33
		FN	0	0	1	6	5	4	8	6	7
✗	FS 5	TP	18	12	14	11	9	11	7	7	9
		FP	70	72	66	26	26	23	14	17	14
		FN	6	12	10	13	15	13	17	17	15
✗	OI 5	TP	14	13	13	11	10	10	10	8	7
		FP	72	73	69	24	26	23	18	18	18
		FN	10	11	11	13	14	14	14	16	17
	FS 10	TP	23	24	24	20	17	17	17	16	15
		FP	104	90	96	49	48	56	30	38	37
		FN	1	0	0	4	7	7	7	8	9
	OI 10	TP	22	23	23	19	18	18	18	18	18
		FP	90	84	76	50	52	47	32	34	29
		FN	2	1	1	5	6	6	6	6	6
✗	FS 10	TP	17	13	13	12	9	10	10	7	8
		FP	70	66	68	31	25	28	14	20	21
		FN	7	11	11	12	15	14	14	17	16
✗	OI 10	TP	17	13	15	11	9	11	10	8	8
		FP	74	66	67	27	27	29	19	19	20
		FN	7	11	9	13	15	13	14	16	16

Table E.7: Post-processed results for logistic regression, patient 37.

Method: Logistic Regression					
KDE	Red		Original	Veto 1	Veto 2
		TP	24	18	16
		FP	93	53	32
		FN	0	6	8

✗		TP	12	8	7
		FP	68	26	18
		FN	12	16	17
	FS 2	TP	22	19	17
		FP	102	48	35
		FN	2	5	7
	OI 2	TP	22	21	19
		FP	113	58	36
		FN	2	3	5
✗	FS 2	TP	15	10	8
		FP	74	22	13
		FN	9	14	16
✗	OI 2	TP	12	7	5
		FP	72	22	15
		FN	12	17	19
	FS 5	TP	23	19	16
		FP	96	56	35
		FN	1	5	8
	OI 5	TP	23	18	16
		FP	110	54	34
		FN	1	6	8
✗	FS 5	TP	14	9	7
		FP	78	21	16
		FN	10	15	17
✗	OI 5	TP	15	11	8
		FP	79	23	17
		FN	9	13	16
	FS 10	TP	24	19	16
		FP	96	49	34
		FN	0	5	8
	OI 10	TP	22	17	16
		FP	100	51	36
		FN	2	7	8
✗	FS 10	TP	12	8	7
		FP	77	23	17
		FN	12	16	17
✗	OI 10	TP	12	8	7
		FP	69	24	18
		FN	12	16	17

Table E.8: Post-processed results SVM, patient 37.

Method: SVM, kernel: rbf, linear, BoxConst: 1, 100														
			Original				Veto 1				Veto 2			
KDE	Red		μ_1	μ_2	μ_3	μ_4	μ_1	μ_2	μ_3	μ_4	μ_1	μ_2	μ_3	μ_4
		TP	24	24	24	24	18	18	17	18	18	18	15	15
		FP	183	181	100	106	55	54	53	53	43	42	35	36
		FN	0	0	0	0	6	6	7	6	6	6	9	9
✗		TP	20	20	13	13	16	16	10	9	12	12	7	7
		FP	102	104	72	65	40	39	23	22	22	24	17	17
		FN	4	4	11	11	8	8	14	15	12	12	17	17
	FS 2	TP	22	24	23	23	21	19	19	20	17	15	16	16
		FP	114	119	116	117	61	52	51	50	44	39	36	36
		FN	2	0	1	1	3	5	5	4	7	9	8	8
	OI 2	TP	24	24	22	22	19	17	21	21	19	18	18	18
		FP	112	107	119	119	51	50	57	56	40	36	34	34
		FN	0	0	2	2	5	7	3	3	5	6	6	6
✗	FS 2	TP	16	16	14	14	12	11	9	9	8	9	8	8
		FP	79	84	69	72	32	28	24	24	20	23	14	15
		FN	8	8	10	10	12	13	15	15	16	15	16	16
✗	OI 2	TP	13	12	13	13	9	7	8	8	7	5	5	5
		FP	75	80	75	75	27	26	23	23	20	22	15	15
		FN	11	12	11	11	15	17	16	16	17	19	19	19

	FS 5	TP	24	24	24	24	21	20	21	20	17	19	18	17
		FP	142	148	107	105	61	59	51	51	39	36	36	39
		FN	0	0	0	0	3	4	3	4	7	5	6	7
	OI 5	TP	24	24	23	23	19	20	18	18	17	17	17	16
		FP	99	113	118	118	49	50	55	55	34	36	35	35
		FN	0	0	1	1	5	4	6	6	7	7	7	8
✗	FS 5	TP	18	20	14	15	15	14	10	11	8	10	8	8
		FP	98	96	79	74	32	36	23	23	22	25	17	18
		FN	6	4	10	9	9	10	14	13	16	14	16	16
✗	OI 5	TP	14	16	16	16	10	10	11	11	8	9	8	8
		FP	70	84	81	82	29	30	24	24	16	16	15	16
		FN	10	8	8	8	14	14	13	13	16	15	16	16
	FS 10	TP	24	24	23	23	19	18	20	19	18	18	16	16
		FP	173	174	102	102	61	58	54	54	41	42	32	35
		FN	0	0	1	1	5	6	4	5	6	6	8	8
	OI 10	TP	24	24	23	22	18	19	19	18	18	17	15	17
		FP	154	153	108	109	57	58	54	53	41	42	36	34
		FN	0	0	1	2	6	5	5	6	6	7	9	7
✗	FS 10	TP	21	21	12	12	16	15	8	9	10	9	6	7
		FP	109	110	72	71	41	39	20	20	26	32	13	14
		FN	3	3	12	12	8	9	16	15	14	15	18	17
✗	OI 10	TP	21	21	14	14	17	16	10	10	11	10	7	7
		FP	103	102	78	74	39	38	26	23	25	26	18	18
		FN	3	3	10	10	7	8	14	14	13	14	17	17

E.2 Cluster balanced training set

Table E.9: Cross-validation results for KNN, patient 37.

Method: KNN, Neighbors: 2, 5, 10								
KDE	Red		μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
		Accuracy	0.88089	0.81069	0.81829	0.016653	0.010064	0.014694
		Specificity	0.92882	0.84702	0.86258	0.019026	0.0082199	0.012668
		Sensitivity	0.18043	0.40492	0.3105	0.038698	0.042492	0.038374
✗		Accuracy	0.89933	0.85153	0.86159	0.012373	0.0076607	0.010725
		Specificity	0.93717	0.88246	0.89597	0.014194	0.0071413	0.0091216
		Sensitivity	0.21627	0.35689	0.30357	0.042209	0.042999	0.033688
	FS 2	Accuracy	0.78651	0.75702	0.78162	0.018848	0.015099	0.010544
		Specificity	0.81535	0.77414	0.80738	0.019402	0.015192	0.013003
		Sensitivity	0.40805	0.5848	0.51798	0.038761	0.042083	0.042217
	OI 2	Accuracy	0.81826	0.76061	0.7621	0.016657	0.015733	0.020755
		Specificity	0.85322	0.78014	0.78426	0.019031	0.013688	0.019934
		Sensitivity	0.33867	0.56766	0.51139	0.039701	0.038332	0.041601
✗	FS 2	Accuracy	0.77206	0.79537	0.81258	0.019804	0.01375	0.012488
		Specificity	0.79744	0.81134	0.83396	0.020881	0.012286	0.012568
		Sensitivity	0.38053	0.57143	0.49237	0.036844	0.037882	0.039903
✗	OI 2	Accuracy	0.85857	0.8111	0.81599	0.013764	0.012124	0.014079
		Specificity	0.89162	0.83649	0.8408	0.014385	0.012919	0.013472
		Sensitivity	0.29238	0.45315	0.44667	0.037974	0.038663	0.034594
	FS 5	Accuracy	0.81922	0.7593	0.77579	0.012125	0.017575	0.015565
		Specificity	0.85318	0.77128	0.79945	0.011991	0.017926	0.01523
		Sensitivity	0.38432	0.61265	0.50512	0.034861	0.035042	0.040165
	OI 5	Accuracy	0.85016	0.78064	0.80901	0.02018	0.018217	0.018975
		Specificity	0.88491	0.80502	0.84639	0.025736	0.014252	0.014886
		Sensitivity	0.27523	0.50969	0.36688	0.038845	0.03613	0.037717
✗	FS 5	Accuracy	0.84882	0.81303	0.83497	0.013937	0.012949	0.011481
		Specificity	0.87736	0.83491	0.86369	0.014708	0.011084	0.010054
		Sensitivity	0.34932	0.53087	0.38979	0.036169	0.040549	0.030917

✗	OI 5	Accuracy	0.88077	0.82548	0.85244	0.014066	0.012879	0.012219
		Specificity	0.91362	0.85035	0.88191	0.01484	0.0094926	0.010002
		Sensitivity	0.2944	0.46781	0.38409	0.035766	0.038799	0.035076
	FS 10	Accuracy	0.8596	0.77932	0.80635	0.014171	0.015302	0.015864
		Specificity	0.8972	0.80424	0.84469	0.016519	0.013193	0.012586
		Sensitivity	0.31522	0.48351	0.39276	0.039325	0.037926	0.036477
	OI 10	Accuracy	0.86527	0.7936	0.80935	0.019677	0.018347	0.018854
		Specificity	0.90432	0.8278	0.85052	0.022329	0.012305	0.014387
		Sensitivity	0.2864	0.43555	0.32146	0.043417	0.039671	0.042915
✗	FS 10	Accuracy	0.88416	0.83288	0.84577	0.0093663	0.010146	0.013419
		Specificity	0.91881	0.85841	0.87974	0.010626	0.0080906	0.0091906
		Sensitivity	0.26349	0.43004	0.32628	0.042222	0.033736	0.038889
✗	OI 10	Accuracy	0.89104	0.83797	0.85542	0.013847	0.01202	0.014039
		Specificity	0.92428	0.86894	0.89093	0.015103	0.0082263	0.0095114
		Sensitivity	0.29006	0.4026	0.32102	0.039392	0.036768	0.043002

Table E.10: Cross-validation results for randf, patient 37.

Method: Random Forest, Trees: 10, 30, 50								
KDE	Red		μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
		Accuracy	0.6817	0.72377	0.74335	0.014264	0.016833	0.018747
		Specificity	0.68049	0.72759	0.75124	0.015807	0.019281	0.021065
		Sensitivity	0.73905	0.74935	0.70879	0.034828	0.032386	0.036581
✗		Accuracy	0.75182	0.78786	0.79409	0.01949	0.014795	0.013394
		Specificity	0.74266	0.78675	0.79357	0.019003	0.013254	0.012022
		Sensitivity	0.71391	0.69071	0.67953	0.022895	0.031829	0.031883
	FS 2	Accuracy	0.69839	0.68563	0.66865	0.019327	0.026533	0.025387
		Specificity	0.70377	0.69563	0.67321	0.020011	0.028027	0.026391
		Sensitivity	0.63967	0.58542	0.61318	0.034613	0.02338	0.034482
	OI 2	Accuracy	0.74528	0.76168	0.76574	0.014066	0.014387	0.013893
		Specificity	0.76267	0.78277	0.78555	0.015634	0.013071	0.013195
		Sensitivity	0.6076	0.58875	0.59336	0.033946	0.035239	0.037137
✗	FS 2	Accuracy	0.78139	0.75109	0.71073	0.016894	0.020366	0.020302
		Specificity	0.78859	0.75092	0.70194	0.016439	0.019353	0.018849
		Sensitivity	0.55667	0.60045	0.58048	0.043133	0.03605	0.029838
✗	OI 2	Accuracy	0.78785	0.79893	0.80223	0.013848	0.012468	0.012155
		Specificity	0.79397	0.81105	0.81331	0.013161	0.010147	0.010153
		Sensitivity	0.55546	0.56628	0.54647	0.031485	0.040228	0.033189
	FS 5	Accuracy	0.65936	0.60445	0.58753	0.022283	0.01853	0.0155
		Specificity	0.65631	0.60909	0.58922	0.02347	0.024077	0.018743
		Sensitivity	0.7142	0.6626	0.63949	0.035075	0.038609	0.039036
	OI 5	Accuracy	0.70103	0.73892	0.74859	0.017837	0.015778	0.012356
		Specificity	0.70765	0.75247	0.76278	0.019654	0.017742	0.013146
		Sensitivity	0.69394	0.63953	0.66055	0.039399	0.040221	0.042284
✗	FS 5	Accuracy	0.76222	0.7112	0.69207	0.020816	0.019392	0.019571
		Specificity	0.75146	0.70122	0.68359	0.019366	0.016967	0.016985
		Sensitivity	0.70853	0.62085	0.58077	0.02966	0.031834	0.034623
✗	OI 5	Accuracy	0.76001	0.78859	0.79513	0.016913	0.013989	0.015601
		Specificity	0.76106	0.79404	0.80399	0.015823	0.01325	0.014456
		Sensitivity	0.63642	0.62247	0.59732	0.033801	0.034075	0.035691
	FS 10	Accuracy	0.69	0.67698	0.69359	0.014914	0.016751	0.015572
		Specificity	0.68965	0.68212	0.70217	0.016998	0.01893	0.019547
		Sensitivity	0.76954	0.67097	0.67955	0.027257	0.037951	0.036192
	OI 10	Accuracy	0.70435	0.72248	0.74066	0.016838	0.01792	0.017191
		Specificity	0.7077	0.73158	0.74882	0.017128	0.017565	0.016004
		Sensitivity	0.71861	0.70764	0.71096	0.026465	0.033284	0.034813
✗	FS 10	Accuracy	0.7371	0.74992	0.75518	0.019327	0.016701	0.015819
		Specificity	0.73107	0.74533	0.75561	0.018969	0.014259	0.014222
		Sensitivity	0.65729	0.64292	0.6075	0.050049	0.031607	0.037286
✗	OI 10	Accuracy	0.76919	0.7941	0.80053	0.017168	0.013661	0.014153
		Specificity	0.77115	0.79786	0.80418	0.015678	0.011326	0.013214
		Sensitivity	0.65911	0.67794	0.67851	0.037401	0.034439	0.028279

Table E.11: Cross-validation results for logistic regression, patient 37.

Method: Logistic Regression				
KDE	Red		μ_1	σ_1
		Accuracy	0.77297	0.012373
		Specificity	0.78879	0.010808
		Sensitivity	0.6045	0.037323
✗		Accuracy	0.82377	0.010667
		Specificity	0.8404	0.010218
		Sensitivity	0.57498	0.039143
	FS 2	Accuracy	0.70105	0.019865
		Specificity	0.71656	0.021253
		Sensitivity	0.54488	0.032417
	OI 2	Accuracy	0.78794	0.014907
		Specificity	0.81412	0.015048
		Sensitivity	0.44562	0.057668
✗	FS 2	Accuracy	0.76727	0.022666
		Specificity	0.78837	0.024603
		Sensitivity	0.42189	0.046197
✗	OI 2	Accuracy	0.77889	0.020517
		Specificity	0.80979	0.020475
		Sensitivity	0.27573	0.041168
	FS 5	Accuracy	0.72922	0.013191
		Specificity	0.74253	0.014669
		Sensitivity	0.56714	0.048958
	OI 5	Accuracy	0.77468	0.014184
		Specificity	0.79646	0.013549
		Sensitivity	0.5048	0.037423
✗	FS 5	Accuracy	0.78174	0.017656
		Specificity	0.80078	0.019095
		Sensitivity	0.5005	0.044945
✗	OI 5	Accuracy	0.82922	0.010407
		Specificity	0.85122	0.010035
		Sensitivity	0.46982	0.036708
	FS 10	Accuracy	0.75566	0.01347
		Specificity	0.77751	0.018085
		Sensitivity	0.55475	0.035253
	OI 10	Accuracy	0.77318	0.013243
		Specificity	0.79294	0.012381
		Sensitivity	0.53668	0.042945
✗	FS 10	Accuracy	0.81376	0.010867
		Specificity	0.83172	0.010638
		Sensitivity	0.5342	0.03932
✗	OI 10	Accuracy	0.83279	0.010341
		Specificity	0.85179	0.01008
		Sensitivity	0.52838	0.046076

Table E.12: Cross-validation results for SVM, patient 37.

Method: SVM, Trees: 10, 30, 50										
KDE	Red		μ_1	μ_2	μ_3	μ_4	σ_1	σ_2	σ_3	σ_4
		Accuracy	0.88771	0.88356	0.77304	0.77571	0.015059	0.014041	0.012069	0.011839
		Specificity	0.96165	0.95636	0.78847	0.79334	0.0038135	0.0039066	0.011276	0.011434
		Sensitivity	0.053638	0.056427	0.60646	0.60363	0.021433	0.018439	0.039021	0.042613
✗		Accuracy	0.91675	0.9127	0.82103	0.8249	0.0097849	0.0094111	0.010541	0.011186
		Specificity	0.97378	0.96889	0.83601	0.84241	0.002627	0.0024749	0.010273	0.011221
		Sensitivity	0.038923	0.054841	0.60104	0.57757	0.018347	0.021181	0.038734	0.041426
	FS 2	Accuracy	0.7411	0.56591	0.91213	0.91012	0.013573	0.02269	0.013657	0.013245
		Specificity	0.77253	0.57161	0.98188	0.97884	0.017482	0.023612	0.0040743	0.0053612
		Sensitivity	0.48689	0.57388	0.092958	0.094817	0.042633	0.041701	0.022828	0.023426
	OI 2	Accuracy	0.71765	0.71917	0.79167	0.78795	0.019232	0.013918	0.019227	0.018609
		Specificity	0.72664	0.73256	0.81585	0.81156	0.019078	0.01274	0.022732	0.021334

		Sensitivity	0.64244	0.60176	0.46688	0.478	0.045452	0.038743	0.061028	0.057705
		Accuracy	0.78344	0.65952	0.86946	0.86489	0.013757	0.018234	0.027298	0.025707
✗	FS 2	Specificity	0.81276	0.67868	0.91314	0.90785	0.01488	0.019426	0.028174	0.026721
		Sensitivity	0.4026	0.39521	0.15983	0.16571	0.039462	0.031796	0.037957	0.03536
		Accuracy	0.80299	0.77462	0.899	0.89643	0.013308	0.014061	0.015821	0.015456
✗	OI 2	Specificity	0.82121	0.78912	0.94718	0.94279	0.013216	0.014494	0.016593	0.016423
		Sensitivity	0.55503	0.56171	0.12219	0.14729	0.046303	0.032727	0.043041	0.044509
		Accuracy	0.82714	0.78769	0.75655	0.75606	0.013451	0.014561	0.013163	0.014604
	FS 5	Specificity	0.88533	0.83374	0.7766	0.77821	0.010706	0.01004	0.012672	0.017877
		Sensitivity	0.21083	0.3152	0.50498	0.53696	0.035469	0.034713	0.049749	0.04664
		Accuracy	0.78983	0.7827	0.77416	0.77421	0.018814	0.015157	0.014151	0.014261
	OI 5	Specificity	0.82028	0.81014	0.79517	0.79503	0.015057	0.01268	0.013739	0.013805
		Sensitivity	0.42523	0.44663	0.50513	0.51599	0.041882	0.040914	0.037242	0.034797
		Accuracy	0.86831	0.82067	0.7968	0.79539	0.010556	0.0108	0.012392	0.014356
✗	FS 5	Specificity	0.91246	0.86044	0.81207	0.81054	0.0074958	0.010926	0.011873	0.013922
		Sensitivity	0.2093	0.25973	0.55642	0.54092	0.034228	0.028544	0.047819	0.045349
		Accuracy	0.84445	0.8222	0.8317	0.83095	0.013087	0.01194	0.011222	0.010563
✗	OI 5	Specificity	0.87331	0.84631	0.85161	0.85092	0.0090411	0.010211	0.010029	0.0097684
		Sensitivity	0.40905	0.44016	0.50345	0.4932	0.043725	0.039082	0.039292	0.039244
		Accuracy	0.87128	0.84868	0.77177	0.77832	0.015162	0.013961	0.011049	0.012912
	FS 10	Specificity	0.94181	0.91217	0.78968	0.80091	0.0062292	0.0057208	0.011287	0.016725
		Sensitivity	0.08671	0.13158	0.53359	0.54359	0.026297	0.02333	0.052831	0.047586
		Accuracy	0.85945	0.84953	0.75472	0.75832	0.020017	0.020364	0.013137	0.013958
	OI 10	Specificity	0.91774	0.90405	0.77195	0.7766	0.010238	0.012684	0.013144	0.013969
		Sensitivity	0.18216	0.19899	0.55291	0.54718	0.034464	0.031785	0.039901	0.040342
		Accuracy	0.90303	0.87841	0.80283	0.8083	0.0097586	0.010195	0.013116	0.012937
✗	FS 10	Specificity	0.95843	0.92819	0.81746	0.8229	0.0046892	0.0060095	0.013735	0.011815
		Sensitivity	0.057715	0.11837	0.59417	0.59875	0.019908	0.023382	0.041637	0.036179
		Accuracy	0.89546	0.88576	0.81886	0.82248	0.012836	0.01258	0.011242	0.011786
✗	OI 10	Specificity	0.94313	0.93224	0.8372	0.84134	0.0054078	0.0055984	0.011278	0.01167
		Sensitivity	0.15924	0.17347	0.53495	0.52563	0.025361	0.024672	0.043871	0.04517

Table E.13: Post-processed results for KNN, patient 37.

Method: KNN, Neighbors: 2, 5, 10										
			Original			Veto 1			Veto 2	
KDE	Red		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
		TP	13	22	21	10	15	13	6	11	9
		FP	36	99	91	15	35	36	3	16	10
		FN	11	2	3	14	9	11	18	13	15
		TP	10	13	12	6	11	9	3	4	6
✗		FP	35	72	61	11	26	23	4	12	8
		FN	14	11	12	18	13	15	21	20	18
		TP	24	24	23	16	16	15	14	15	11
	FS 2	FP	107	110	107	37	41	44	19	29	25
		FN	0	0	1	8	8	9	10	9	13
		TP	22	24	24	19	15	17	9	13	13
	OI 2	FP	87	114	108	38	38	44	20	31	28
		FN	2	0	0	5	9	7	15	11	11
		TP	17	16	16	11	8	10	7	8	10
✗	FS 2	FP	93	95	82	29	31	25	19	20	22
		FN	7	8	8	13	16	14	17	16	14
		TP	13	15	15	7	8	10	3	5	5
✗	OI 2	FP	69	80	80	19	27	31	12	16	20
		FN	11	9	9	17	16	14	21	19	19
		TP	23	24	23	15	18	16	14	18	14
	FS 5	FP	86	110	114	38	48	42	17	32	24
		FN	1	0	1	9	6	8	10	6	10
		TP	19	24	21	13	16	15	5	15	12
	OI 5	FP	65	103	92	31	45	40	16	26	18
		FN	5	0	3	11	8	9	19	9	12
		TP	13	14	14	10	9	9	5	7	7
✗	FS 5	FP	68	76	72	25	30	28	14	15	17

		FN	11	10	10	14	15	15	19	17	17
		TP	12	14	13	7	11	9	5	6	5
✗	OI 5	FP	47	78	67	16	39	30	8	20	15
		FN	12	10	11	17	13	15	19	18	19
		TP	21	24	23	14	15	15	11	12	12
	FS 10	FP	65	101	93	25	41	41	12	19	9
		FN	3	0	1	10	9	9	13	12	12
		TP	17	23	22	14	18	15	8	13	7
	OI 10	FP	45	101	93	27	32	32	8	12	6
		FN	7	1	2	10	6	9	16	11	17
		TP	11	16	13	6	12	7	4	7	6
✗	FS 10	FP	56	73	65	25	26	23	10	15	10
		FN	13	8	11	18	12	17	20	17	18
		TP	12	13	10	7	9	7	4	7	3
✗	OI 10	FP	44	70	64	21	27	25	7	15	9
		FN	12	11	14	17	15	17	20	17	21

Table E.14: Post-processed results for random forest, patient 37.

Method: Random Forest, Trees: 10, 30, 50											
			Original			Veto 1			Veto 2		
KDE	Red		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
		TP	24	24	24	15	14	15	17	17	17
		FP	136	118	118	39	30	44	34	30	29
		FN	0	0	0	9	10	9	7	7	7
		TP	23	23	23	16	17	16	17	18	17
✗		FP	96	85	85	35	36	31	31	28	27
		FN	1	1	1	8	7	8	7	6	7
		TP	24	24	24	16	14	13	14	15	16
	FS 2	FP	136	132	137	49	34	37	34	34	28
		FN	0	0	0	8	10	11	10	9	8
		TP	24	24	24	16	15	18	17	17	16
	OI 2	FP	112	107	108	46	47	40	33	31	29
		FN	0	0	0	8	9	6	7	7	8
		TP	22	23	22	17	13	13	15	16	14
✗	FS 2	FP	93	92	97	41	32	33	28	28	32
		FN	2	1	2	7	11	11	9	8	10
		TP	23	23	23	15	15	14	17	16	17
✗	OI 2	FP	82	83	82	39	37	39	29	29	27
		FN	1	1	1	9	9	10	7	8	7
		TP	24	24	24	12	15	14	15	15	16
	FS 5	FP	142	161	166	32	22	24	29	32	29
		FN	0	0	0	12	9	10	9	9	8
		TP	24	24	24	14	16	15	16	14	14
	OI 5	FP	121	116	117	36	38	43	27	29	27
		FN	0	0	0	10	8	9	8	10	10
		TP	23	23	23	15	13	13	18	16	15
✗	FS 5	FP	92	100	106	41	29	18	31	25	28
		FN	1	1	1	9	11	11	6	8	9
		TP	23	23	23	13	16	16	16	15	13
✗	OI 5	FP	87	85	82	41	40	32	34	29	30
		FN	1	1	1	11	8	8	8	9	11
		TP	24	24	24	15	11	14	17	16	15
	FS 10	FP	138	131	129	36	30	30	32	29	31
		FN	0	0	0	9	13	10	7	8	9
		TP	24	24	24	15	12	15	16	17	18
	OI 10	FP	129	119	114	40	36	41	34	30	31
		FN	0	0	0	9	12	9	8	7	6
		TP	23	23	23	13	14	14	16	16	14
✗	FS 10	FP	92	94	91	28	29	33	27	24	30
		FN	1	1	1	11	10	10	8	8	10
		TP	23	23	23	17	14	16	15	19	16
✗	OI 10	FP	83	83	82	32	39	42	33	26	27
		FN	1	1	1	7	10	8	9	5	8

Table E.15: Post-processed results for logistic regression, patient 37.

Method: Logistic Regression					
KDE	Red		Original	Veto 1	Veto 2
		TP	24	15	15
		FP	110	43	25
		FN	0	9	9
✗		TP	15	9	6
		FP	75	25	16
		FN	9	15	18
	FS 2	TP	24	15	15
		FP	139	37	30
		FN	0	9	9
	OI 2	TP	19	16	16
		FP	111	38	27
		FN	5	8	8
✗	FS 2	TP	14	9	7
		FP	87	30	26
		FN	10	15	17
✗	OI 2	TP	16	11	6
		FP	75	22	14
		FN	8	13	18
	FS 5	TP	22	14	16
		FP	130	33	30
		FN	2	10	8
	OI 5	TP	22	17	15
		FP	108	39	29
		FN	2	7	9
✗	FS 5	TP	14	9	7
		FP	87	32	19
		FN	10	15	17
✗	OI 5	TP	12	9	6
		FP	81	23	17
		FN	12	15	18
	FS 10	TP	24	16	14
		FP	115	38	24
		FN	0	8	10
	OI 10	TP	23	17	15
		FP	111	45	28
		FN	1	7	9
✗	FS 10	TP	13	11	8
		FP	82	25	19
		FN	11	13	16
✗	OI 10	TP	13	9	5
		FP	77	30	17
		FN	11	15	19

Table E.16: Post-processed results SVM, patient 37.

Method: SVM, kernel: rbf, linear, BoxConst: 1, 100														
			Original				Veto 1				Veto 2			
KDE	Red		μ_1	μ_2	μ_3	μ_4	μ_1	μ_2	μ_3	μ_4	μ_1	μ_2	μ_3	μ_4
		TP	8	9	23	23	2	2	16	14	0	0	15	16
		FP	34	33	109	105	1	4	44	42	0	0	27	24
		FN	16	15	1	1	22	22	8	10	24	24	9	8
✗		TP	3	4	15	14	1	1	10	8	0	0	6	7
		FP	16	23	79	78	5	7	29	29	0	0	16	16
		FN	21	20	9	10	23	23	14	16	24	24	18	17
	FS 2	TP	23	23	9	9	12	13	7	8	10	13	6	6
		FP	109	181	10	11	39	45	4	4	25	32	1	1
		FN	1	1	15	15	12	11	17	16	14	11	18	18

	OI 2	TP	24	24	19	20	15	13	17	17	15	16	13	14
		FP	117	124	96	98	39	41	38	37	35	35	26	26
		FN	0	0	5	4	9	11	7	7	9	8	11	10
✗	FS 2	TP	13	22	6	7	9	13	5	6	7	14	3	4
		FP	87	113	34	34	25	33	8	5	19	25	7	5
		FN	11	2	18	17	15	11	19	18	17	10	21	20
✗	OI 2	TP	15	17	2	5	9	12	2	4	9	11	1	2
		FP	77	93	28	31	25	28	8	12	22	17	3	4
		FN	9	7	22	19	15	12	22	20	15	13	23	22
	FS 5	TP	18	23	23	23	9	16	16	16	2	7	15	15
		FP	84	111	116	121	36	40	41	43	11	19	28	31
		FN	6	1	1	1	15	8	8	8	22	17	9	9
	OI 5	TP	21	21	23	23	14	14	17	17	11	13	15	13
		FP	93	104	103	104	43	40	44	45	25	24	28	30
		FN	3	3	1	1	10	10	7	7	13	11	9	11
✗	FS 5	TP	11	15	15	12	5	11	10	8	5	6	8	5
		FP	60	80	91	88	25	28	28	26	6	12	18	19
		FN	13	9	9	12	19	13	14	16	19	18	16	19
✗	OI 5	TP	11	14	14	13	8	10	11	10	6	6	7	6
		FP	69	78	83	83	27	28	23	24	12	20	14	14
		FN	13	10	10	11	16	14	13	14	18	18	17	18
	FS 10	TP	9	14	22	23	2	6	17	13	0	0	15	14
		FP	55	71	110	108	11	24	43	44	0	4	26	28
		FN	15	10	2	1	22	18	7	11	24	24	9	10
	OI 10	TP	16	15	24	24	6	7	18	18	2	3	15	15
		FP	62	61	118	113	18	21	47	46	0	2	31	32
		FN	8	9	0	0	18	17	6	6	22	21	9	9
✗	FS 10	TP	7	10	16	15	4	6	9	8	2	3	7	7
		FP	34	53	84	81	11	19	25	26	2	4	20	19
		FN	17	14	8	9	20	18	15	16	22	21	17	17
✗	OI 10	TP	10	10	14	14	8	7	9	9	1	3	6	6
		FP	32	42	80	78	11	16	33	33	3	4	19	18
		FN	14	14	10	10	16	17	15	15	23	21	18	18

F

Appendix 6: HMS Patient 39 result

F.1 Randomly balanced training set

Table F.1: Cross-validation results for KNN, patient 39.

Method: KNN, Neighbors: 2, 5, 10								
KDE	Red		μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
		Accuracy	0.92972	0.97938	0.98497	0.037229	0.009124	0.00083187
		Specificity	0.93263	0.9836	0.9899	0.038037	0.010929	0.0025688
		Sensitivity	0.63034	0.59615	0.52991	0.09188	0.096154	0.08547
✗		Accuracy	0.93871	0.9883	0.99364	0.050383	0.0015461	0.0022272
		Specificity	0.94022	0.99101	0.99637	0.052607	0.0018145	0.0024953
		Sensitivity	0.61538	0.49786	0.49786	0.38462	0.11325	0.11325
	FS 2	Accuracy	0.91845	0.95393	0.9622	0.029336	0.033902	0.026977
		Specificity	0.92227	0.95827	0.96696	0.02973	0.036938	0.030298
		Sensitivity	0.53632	0.57906	0.55128	0.074786	0.19017	0.21795
	OI 2	Accuracy	0.83038	0.85937	0.87025	0.10286	0.036504	0.0061536
		Specificity	0.83212	0.86169	0.87194	0.10522	0.037692	0.0047262
		Sensitivity	0.61966	0.65171	0.72863	0.15812	0.040598	0.11752
✗	FS 2	Accuracy	0.96003	0.95219	0.9659	0.029821	0.039163	0.022448
		Specificity	0.96201	0.95335	0.96735	0.030813	0.03909	0.022827
		Sensitivity	0.5812	0.74573	0.69658	0.19658	0.023504	0.081197
✗	OI 2	Accuracy	0.9012	0.93769	0.94326	0.068346	0.016071	0.041327
		Specificity	0.90265	0.93899	0.94459	0.070153	0.016435	0.042569
		Sensitivity	0.59829	0.69658	0.67521	0.2906	0.081197	0.21368
	FS 5	Accuracy	0.89575	0.95553	0.9794	0.0021575	0.027569	0.01113
		Specificity	0.89771	0.95987	0.98431	0.0030431	0.029867	0.01364
		Sensitivity	0.71795	0.56838	0.5406	0.051282	0.12393	0.15171
	OI 5	Accuracy	0.8899	0.88435	0.88574	0.071104	0.028432	0.0040492
		Specificity	0.893	0.88658	0.88794	0.072331	0.029914	0.004603
		Sensitivity	0.5641	0.69017	0.67949	0.10256	0.07906	0.012821
✗	FS 5	Accuracy	0.92937	0.95807	0.98225	0.065371	0.032156	0.0087314
		Specificity	0.93023	0.9591	0.98436	0.066753	0.032588	0.0092168
		Sensitivity	0.74145	0.76282	0.59188	0.20299	0.070513	0.13034
✗	OI 5	Accuracy	0.94874	0.94443	0.94733	0.040361	0.021737	0.035376
		Specificity	0.95028	0.94576	0.94887	0.041414	0.022129	0.03677
		Sensitivity	0.64744	0.69658	0.63675	0.1859	0.081197	0.25214
	FS 10	Accuracy	0.92216	0.95575	0.98436	0.055363	0.01788	0.0061726
		Specificity	0.92429	0.95939	0.9893	0.055708	0.019391	0.0079644
		Sensitivity	0.70726	0.62393	0.52991	0.014957	0.068376	0.08547
	OI 10	Accuracy	0.92042	0.96088	0.96982	0.061733	0.02763	0.020713
		Specificity	0.92286	0.9646	0.97497	0.061806	0.029929	0.022973
		Sensitivity	0.67949	0.63462	0.50214	0.012821	0.13462	0.11325
✗	FS 10	Accuracy	0.9326	0.96653	0.9755	0.062517	0.025947	0.017733
		Specificity	0.93348	0.96725	0.97684	0.063879	0.026328	0.018627
		Sensitivity	0.74145	0.82906	0.71368	0.20299	0.059829	0.17521
✗	OI 10	Accuracy	0.94824	0.98221	0.99341	0.042733	0.00025377	0.00057158
		Specificity	0.94904	0.98394	0.99576	0.043781	0.00056634	8.3559e-05

		Sensitivity	0.77991	0.67949	0.58547	0.16453	0.012821	0.029915

Table F.2: Cross-validation results for randf, patient 39.

Method: Random Forest, Trees: 10, 30, 50								
KDE	Red		μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
		Accuracy	0.81828	0.86506	0.86286	0.038376	0.066648	0.07899
		Specificity	0.82018	0.86715	0.86529	0.039509	0.067821	0.079955
		Sensitivity	0.65171	0.67949	0.64103	0.040598	0.012821	0.025641
		Accuracy	0.8982	0.88785	0.90321	0.031287	0.024931	0.023108
✗		Specificity	0.8995	0.88756	0.90375	0.030973	0.02517	0.023342
		Sensitivity	0.65812	0.93376	0.80128	0.11966	0.010684	0.032051
		Accuracy	0.81756	0.81273	0.84467	0.033689	0.049342	0.066082
	FS 2	Specificity	0.81909	0.8139	0.84723	0.033754	0.049896	0.067897
		Sensitivity	0.6688	0.70726	0.62393	0.053419	0.014957	0.068376
		Accuracy	0.82822	0.80199	0.78387	0.00072171	0.045878	0.039656
	OI 2	Specificity	0.82979	0.80305	0.78473	0.00045996	0.047051	0.040731
		Sensitivity	0.6688	0.71795	0.71795	0.053419	0.051282	0.051282
		Accuracy	0.86806	0.90336	0.90676	0.025545	0.021828	0.018428
✗	FS 2	Specificity	0.86933	0.90464	0.90824	0.026027	0.022072	0.01847
		Sensitivity	0.63034	0.6688	0.64103	0.09188	0.053419	0.025641
		Accuracy	0.88707	0.89629	0.90583	0.037746	0.040173	0.047558
✗	OI 2	Specificity	0.88751	0.89734	0.90656	0.038445	0.041076	0.048487
		Sensitivity	0.7906	0.6859	0.75214	0.098291	0.14744	0.13675
		Accuracy	0.81394	0.8397	0.83044	0.067739	0.069702	0.078287
	FS 5	Specificity	0.8155	0.84187	0.83253	0.06883	0.071203	0.079861
		Sensitivity	0.67949	0.65171	0.65171	0.012821	0.040598	0.040598
		Accuracy	0.8345	0.81774	0.82263	0.053266	0.019988	0.067837
	OI 5	Specificity	0.83591	0.81892	0.82395	0.054576	0.020918	0.069279
		Sensitivity	0.71795	0.71795	0.71795	0.051282	0.051282	0.051282
		Accuracy	0.88578	0.91022	0.90758	0.0078281	0.025114	0.022115
✗	FS 5	Specificity	0.88641	0.9108	0.90833	0.0078088	0.025354	0.02216
		Sensitivity	0.7735	0.80128	0.7735	0.0042735	0.032051	0.0042735
		Accuracy	0.89432	0.89302	0.89228	0.027107	0.033675	0.033657
✗	OI 5	Specificity	0.89482	0.89386	0.89275	0.028108	0.033986	0.03396
		Sensitivity	0.77991	0.73504	0.80128	0.16453	0.042735	0.032051
		Accuracy	0.76284	0.83943	0.82071	0.041757	0.075381	0.046091
	FS 10	Specificity	0.76451	0.84162	0.82197	0.043842	0.077617	0.0473
		Sensitivity	0.63462	0.66239	0.71795	0.13462	0.10684	0.051282
		Accuracy	0.86013	0.81652	0.83853	0.073611	0.089493	0.085064
	OI 10	Specificity	0.86251	0.81747	0.84004	0.074515	0.090808	0.086039
		Sensitivity	0.64103	0.74573	0.70726	0.025641	0.023504	0.014957
		Accuracy	0.8888	0.91479	0.91775	0.048423	0.054012	0.055185
✗	FS 10	Specificity	0.88944	0.91591	0.9189	0.04974	0.054242	0.055788
		Sensitivity	0.74145	0.70726	0.69658	0.20299	0.014957	0.081197
		Accuracy	0.89861	0.90656	0.90701	0.03748	0.047576	0.04337
✗	OI 10	Specificity	0.89856	0.90747	0.90738	0.037971	0.047956	0.044265
		Sensitivity	0.8953	0.73504	0.81838	0.049145	0.042735	0.12607

Table F.3: Cross-validation results for logistic regression, patient 39.

Method: Logistic Regression				
KDE	Red		μ_1	σ_1
		Accuracy	0.71209	0.15969
		Specificity	0.71759	0.15907
		Sensitivity	0.18162	0.12607
		Accuracy	0.81956	0.023318
✗		Specificity	0.82241	0.023771
		Sensitivity	0.32051	0.012821
		Accuracy	0.81603	0.10486
	FS 2	Specificity	0.8194	0.10873

		Sensitivity	0.55128	0.21795
	OI 2	Accuracy	0.84642	0.044044
		Specificity	0.84795	0.045279
		Sensitivity	0.71795	0.051282
✗	FS 2	Accuracy	0.90924	0.044951
		Specificity	0.91133	0.045443
		Sensitivity	0.54701	0.008547
✗	OI 2	Accuracy	0.94362	0.0037486
		Specificity	0.94476	0.0034804
		Sensitivity	0.74573	0.023504
	FS 5	Accuracy	0.80426	0.0693
		Specificity	0.80608	0.072083
		Sensitivity	0.67308	0.17308
	OI 5	Accuracy	0.87043	0.038969
		Specificity	0.87254	0.040537
		Sensitivity	0.69017	0.07906
✗	FS 5	Accuracy	0.89554	0.035387
		Specificity	0.89718	0.035453
		Sensitivity	0.60256	0.064103
✗	OI 5	Accuracy	0.94198	0.018294
		Specificity	0.9446	0.01902
		Sensitivity	0.49145	0.047009
	FS 10	Accuracy	0.70169	0.045836
		Specificity	0.70426	0.043071
		Sensitivity	0.49573	0.2735
	OI 10	Accuracy	0.8547	0.037119
		Specificity	0.85764	0.037639
		Sensitivity	0.57479	0.036325
✗	FS 10	Accuracy	0.84867	0.040022
		Specificity	0.84969	0.039718
		Sensitivity	0.65812	0.11966
✗	OI 10	Accuracy	0.88688	0.022523
		Specificity	0.88806	0.022784
		Sensitivity	0.6688	0.053419

Table F.4: Cross-validation results for SVM, patient 39.

Method: SVM, Trees: 10, 30, 50										
KDE	Red		μ_1	μ_2	μ_3	μ_4	σ_1	σ_2	σ_3	σ_4
		Accuracy	0.98962	0.98929	0.89609	0.89609	0.0017891	0.0014586	0.061988	0.061988
		Specificity	1	0.99967	0.89953	0.89953	0	0.00033333	0.064195	0.064195
		Sensitivity	0	0	0.59615	0.59615	0	0	0.096154	0.096154
✗		Accuracy	0.99434	0.99434	0.93235	0.93235	0.00077758	0.00077758	0.022198	0.022198
		Specificity	1	1	0.93401	0.93401	0	0	0.022407	0.022407
		Sensitivity	0	0	0.64103	0.64103	0	0	0.025641	0.025641
	FS 2	Accuracy	0.68217	0.80882	0.91428	0.83364	0.060225	0.087734	0.014343	0.066291
		Specificity	0.68099	0.81003	0.91742	0.83643	0.061654	0.089359	0.013246	0.06843
		Sensitivity	0.82265	0.71795	0.62393	0.59615	0.10043	0.051282	0.068376	0.096154
	OI 2	Accuracy	0.82386	0.81722	0.84488	0.8382	0.087571	0.033351	0.0030996	0.0031755
		Specificity	0.82591	0.8191	0.84665	0.84058	0.089906	0.034429	0.0020198	0.0014244
		Sensitivity	0.66239	0.65171	0.69017	0.63462	0.10684	0.040598	0.07906	0.13462
✗	FS 2	Accuracy	0.93413	0.94646	0.94437	0.91177	0.024354	0.0073659	0.012018	0.046356
		Specificity	0.93707	0.95003	0.94697	0.91406	0.025469	0.0078272	0.011093	0.047038
		Sensitivity	0.4359	0.32051	0.50214	0.51923	0.10256	0.012821	0.11325	0.019231
✗	OI 2	Accuracy	0.89641	0.88052	0.94476	0.95082	0.029908	0.002565	0.0052434	0.0059425
		Specificity	0.89711	0.88058	0.94554	0.9522	0.031113	0.0030601	0.005343	0.0062391
		Sensitivity	0.74145	0.85684	0.80128	0.69658	0.20299	0.087607	0.032051	0.081197
	FS 5	Accuracy	0.68393	0.71495	0.8986	0.82556	0.30389	0.26815	0.017564	0.090598
		Specificity	0.68733	0.71758	0.90163	0.82763	0.31267	0.27558	0.018968	0.092968
		Sensitivity	0.5	0.58333	0.62393	0.66239	0.5	0.41667	0.068376	0.10684
	OI 5	Accuracy	0.74537	0.75348	0.87336	0.84986	0.23975	0.21812	0.031983	0.052773
		Specificity	0.7493	0.75779	0.87583	0.85279	0.24796	0.22579	0.033825	0.054123
		Sensitivity	0.5	0.46154	0.66239	0.58547	0.5	0.46154	0.10684	0.029915

✗	FS 5	Accuracy	0.98855	0.98149	0.92581	0.92947	0.0032149	0.0070554	0.012274	0.027219
		Specificity	0.99363	0.98635	0.92741	0.93186	0.0042083	0.0078851	0.012791	0.027437
		Sensitivity	0.1047	0.13248	0.65171	0.50855	0.049145	0.021368	0.040598	0.047009
✗	OI 5	Accuracy	0.97773	0.96997	0.9426	0.92624	0.0042201	0.0014623	0.008011	0.031882
		Specificity	0.9833	0.97531	0.94466	0.92822	0.0034752	0.00088792	0.0085079	0.03324
		Sensitivity	0	0.027778	0.58547	0.60684	0	0.027778	0.029915	0.16239
	FS 10	Accuracy	0.64527	0.66146	0.86085	0.87077	0.34256	0.3257	0.049221	0.039307
		Specificity	0.64833	0.66466	0.86391	0.87424	0.35167	0.33466	0.051906	0.041572
		Sensitivity	0.5	0.5	0.60684	0.56838	0.5	0.5	0.16239	0.12393
	OI 10	Accuracy	0.98664	0.98499	0.94145	0.91204	0.0011851	0.0028375	0.0050102	0.024402
		Specificity	0.997	0.99533	0.94457	0.91557	0.003	0.0046667	0.0034319	0.025568
		Sensitivity	0	0	0.66239	0.58547	0	0	0.10684	0.029915
✗	FS 10	Accuracy	0.99415	0.99358	0.91912	0.91979	0.00058961	2.57e-05	0.015738	0.025806
		Specificity	0.99981	0.99924	0.92014	0.92065	0.00018889	0.00075557	0.016098	0.025675
		Sensitivity	0	0	0.74573	0.76282	0	0	0.023504	0.070513
✗	OI 10	Accuracy	0.99434	0.99396	0.9467	0.91512	0.00077758	0.00040164	0.0021741	0.029405
		Specificity	1	0.99962	0.94824	0.9165	0	0.00037779	0.0022723	0.029462
		Sensitivity	0	0	0.6688	0.6688	0	0	0.053419	0.053419

Table F.5: Post-processed results for KNN, patient 39.

Method: KNN, Neighbors: 2, 5, 10											
			Original			Veto 1			Veto 2		
KDE	Red		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
		TP	2	2	2	2	2	2	2	2	2
		FP	42	10	6	36	8	4	22	5	4
		FN	0	0	0	0	0	0	0	0	0
✗		TP	1	0	0	1	0	0	1	0	0
		FP	11	3	1	1	1	0	0	1	0
		FN	1	2	2	1	2	2	1	2	2
	FS 2	TP	2	2	2	2	2	2	2	2	2
		FP	43	18	18	38	14	14	24	12	8
		FN	0	0	0	0	0	0	0	0	0
	OI 2	TP	2	2	2	2	2	2	2	2	2
		FP	58	48	47	44	43	41	33	33	36
		FN	0	0	0	0	0	0	0	0	0
✗	FS 2	TP	0	1	0	0	1	0	0	1	0
		FP	12	12	10	3	3	4	2	3	4
		FN	2	1	2	2	1	2	2	1	2
✗	OI 2	TP	1	1	1	1	1	1	1	0	0
		FP	12	10	8	2	4	1	1	4	1
		FN	1	1	1	1	1	1	1	2	2
	FS 5	TP	2	2	2	2	2	2	2	2	2
		FP	44	23	10	39	20	7	29	12	6
		FN	0	0	0	0	0	0	0	0	0
	OI 5	TP	2	2	2	2	2	2	2	2	2
		FP	43	47	46	40	42	42	31	28	36
		FN	0	0	0	0	0	0	0	0	0
✗	FS 5	TP	1	1	0	1	1	0	1	0	0
		FP	14	9	5	6	3	1	2	3	1
		FN	1	1	2	1	1	2	1	2	2
✗	OI 5	TP	1	0	1	1	0	1	1	0	0
		FP	10	12	10	1	3	1	1	1	1
		FN	1	2	1	1	2	1	1	2	2
	FS 10	TP	2	2	2	2	2	2	2	2	2
		FP	41	24	6	38	20	4	25	11	5
		FN	0	0	0	0	0	0	0	0	0
	OI 10	TP	2	2	2	2	2	2	2	2	2
		FP	36	16	13	34	13	9	27	7	7
		FN	0	0	0	0	0	0	0	0	0
✗	FS 10	TP	1	0	1	1	0	1	1	0	0
		FP	14	9	5	6	3	1	0	3	1
		FN	1	2	1	1	2	1	1	2	2

✗	OI 10	TP	1	0	0	1	0	0	0	0	0
		FP	11	5	4	3	3	3	1	1	1
		FN	1	2	2	1	2	2	2	2	2

Table F.6: Post-processed results for random forest, patient 39.

Method: Random Forest, Trees: 10, 30, 50											
			Original			Veto 1			Veto 2		
KDE	Red		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
		TP	2	2	2	2	2	2	2	2	2
		FP	58	42	41	49	35	36	40	33	34
		FN	0	0	0	0	0	0	0	0	0
✗		TP	2	1	1	2	1	1	2	1	1
		FP	14	18	16	5	8	7	4	4	4
		FN	0	1	1	0	1	1	0	1	1
	FS 2	TP	2	2	2	2	2	2	2	2	2
		FP	59	55	47	47	49	42	37	40	36
		FN	0	0	0	0	0	0	0	0	0
	OI 2	TP	2	2	2	2	2	2	2	2	2
		FP	56	62	67	49	50	54	39	41	46
		FN	0	0	0	0	0	0	0	0	0
✗	FS 2	TP	1	1	1	1	1	1	1	1	1
		FP	18	16	15	9	8	5	4	1	2
		FN	1	1	1	1	1	1	1	1	1
✗	OI 2	TP	1	1	1	1	1	1	1	1	1
		FP	17	18	17	8	8	7	4	4	1
		FN	1	1	1	1	1	1	1	1	1
	FS 5	TP	2	2	2	2	2	2	2	2	2
		FP	54	47	47	47	38	41	38	33	36
		FN	0	0	0	0	0	0	0	0	0
	OI 5	TP	2	2	2	2	2	2	2	2	2
		FP	51	61	51	42	51	41	36	42	33
		FN	0	0	0	0	0	0	0	0	0
✗	FS 5	TP	1	1	1	1	1	1	1	1	1
		FP	17	17	16	9	8	8	5	4	4
		FN	1	1	1	1	1	1	1	1	1
✗	OI 5	TP	1	1	1	1	1	1	1	1	1
		FP	17	18	17	8	8	8	4	2	4
		FN	1	1	1	1	1	1	1	1	1
	FS 10	TP	2	2	2	2	2	2	2	2	2
		FP	67	49	55	50	41	48	41	33	40
		FN	0	0	0	0	0	0	0	0	0
	OI 10	TP	2	2	2	2	2	2	2	2	2
		FP	40	52	47	36	42	39	31	36	38
		FN	0	0	0	0	0	0	0	0	0
✗	FS 10	TP	1	1	1	1	1	1	1	1	1
		FP	17	17	17	8	7	8	4	1	3
		FN	1	1	1	1	1	1	1	1	1
✗	OI 10	TP	1	1	1	1	1	1	1	1	1
		FP	17	17	17	8	8	8	4	2	1
		FN	1	1	1	1	1	1	1	1	1

Table F.7: Post-processed results for logistic regression, patient 39.

Method: Logistic Regression					
KDE	Red		Original	Veto 1	Veto 2
		TP	1	1	1
		FP	82	58	22
		FN	1	1	1

✗		TP	1	1	1
		FP	21	6	5
		FN	1	1	1
	FS 2	TP	2	2	2
		FP	51	34	27
		FN	0	0	0
	OI 2	TP	2	2	2
		FP	51	41	35
		FN	0	0	0
✗	FS 2	TP	1	1	0
		FP	9	3	2
		FN	1	1	2
✗	OI 2	TP	1	1	0
		FP	11	4	3
		FN	1	1	2
	FS 5	TP	2	2	2
		FP	67	51	35
		FN	0	0	0
	OI 5	TP	2	2	2
		FP	47	35	34
		FN	0	0	0
✗	FS 5	TP	1	1	1
		FP	14	3	2
		FN	1	1	1
✗	OI 5	TP	0	0	0
		FP	8	5	2
		FN	2	2	2
	FS 10	TP	2	2	1
		FP	89	54	28
		FN	0	0	1
	OI 10	TP	2	2	2
		FP	51	42	29
		FN	0	0	0
✗	FS 10	TP	2	2	1
		FP	18	6	5
		FN	0	0	1
✗	OI 10	TP	1	1	1
		FP	15	5	3
		FN	1	1	1

Table F.8: Post-processed results SVM, patient 39.

Method: SVM, kernel: rbf, linear, BoxConst: 1, 100														
			Original				Veto 1				Veto 2			
KDE	Red		μ_1	μ_2	μ_3	μ_4	μ_1	μ_2	μ_3	μ_4	μ_1	μ_2	μ_3	μ_4
		TP	0	0	2	2	0	0	2	2	0	0	2	2
		FP	0	0	32	32	0	0	26	26	0	0	20	20
		FN	2	2	0	0	2	2	0	0	2	2	0	0
✗		TP	0	0	0	0	0	0	0	0	0	0	0	0
		FP	0	0	9	9	0	0	5	5	0	0	4	4
		FN	2	2	2	2	2	2	2	2	2	2	2	2
	FS 2	TP	2	2	2	2	2	2	2	2	2	2	2	2
		FP	80	61	44	55	59	40	38	41	49	26	31	34
		FN	0	0	0	0	0	0	0	0	0	0	0	0
	OI 2	TP	2	2	2	2	2	2	2	2	2	2	2	2
		FP	55	60	59	62	44	50	49	48	35	38	41	43
		FN	0	0	0	0	0	0	0	0	0	0	0	0
✗	FS 2	TP	0	0	0	0	0	0	0	0	0	0	0	0
		FP	10	8	12	9	5	2	2	5	1	1	2	4
		FN	2	2	2	2	2	2	2	2	2	2	2	2
✗	OI 2	TP	1	1	1	0	1	1	1	0	1	1	0	0
		FP	15	16	10	10	8	8	4	4	2	4	4	3
		FN	1	1	1	2	1	1	1	2	1	1	2	2

	FS 5	TP	1	2	2	2	1	2	2	2	1	2	2	2
		FP	46	45	40	50	19	21	35	39	20	20	28	31
		FN	1	0	0	0	1	0	0	0	1	0	0	0
	OI 5	TP	1	1	2	2	1	1	2	2	1	1	2	2
		FP	42	45	45	47	20	22	38	43	18	19	35	38
		FN	1	1	0	0	1	1	0	0	1	1	0	0
✗	FS 5	TP	0	0	0	0	0	0	0	0	0	0	0	0
		FP	2	2	12	12	2	2	5	5	1	1	2	4
		FN	2	2	2	2	2	2	2	2	2	2	2	2
✗	OI 5	TP	0	0	0	1	0	0	0	1	0	0	0	0
		FP	3	4	9	10	1	1	4	5	0	0	4	4
		FN	2	2	2	1	2	2	2	1	2	2	2	2
	FS 10	TP	1	1	2	2	1	1	2	2	1	1	2	2
		FP	46	46	46	45	18	19	40	38	20	20	31	30
		FN	1	1	0	0	1	1	0	0	1	1	0	0
	OI 10	TP	0	0	2	2	0	0	2	2	0	0	2	2
		FP	1	2	32	38	1	2	28	32	1	1	23	28
		FN	2	2	0	0	2	2	0	0	2	2	0	0
✗	FS 10	TP	0	0	1	1	0	0	1	1	0	0	1	1
		FP	0	0	14	11	0	0	5	5	0	0	5	3
		FN	2	2	1	1	2	2	1	1	2	2	1	1
✗	OI 10	TP	0	0	0	1	0	0	0	0	0	0	0	0
		FP	0	0	10	12	0	0	5	5	0	0	3	2
		FN	2	2	2	1	2	2	2	2	2	2	2	2

F.2 Cluster balanced training set

Table F.9: Cross-validation results for KNN, patient 39.

Method: KNN, Neighbors: 2, 5, 10								
KDE	Red		μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
		Accuracy	0.97305	0.99162	0.99096	0.019023	0.0017663	0.0017739
		Specificity	0.97845	0.99864	0.99863	0.020885	0.00069336	0.0013689
		Sensitivity	0.41026	0.32051	0.24359	0.25641	0.012821	0.089744
✗		Accuracy	0.97821	0.99524	0.99507	0.01615	0.00024973	0.00079561
		Specificity	0.98088	0.99944	0.99928	0.017985	0.00019767	0.00072046
		Sensitivity	0.45513	0.25427	0.24359	0.37821	0.023504	0.089744
	FS 2	Accuracy	0.95109	0.86607	0.98359	0.038992	0.07973	0.002526
		Specificity	0.9549	0.8706	0.98779	0.03977	0.080732	0.0042136
		Sensitivity	0.5641	0.40385	0.54274	0.10256	0.096154	0.23504
	OI 2	Accuracy	0.91978	0.96999	0.96822	0.059735	0.0063435	0.0027007
		Specificity	0.92354	0.97544	0.97328	0.061122	0.0067687	0.0027159
		Sensitivity	0.52564	0.44231	0.47009	0.14103	0.057692	0.08547
✗	FS 2	Accuracy	0.94669	0.99392	0.99039	0.046545	0.0010661	0.0043477
		Specificity	0.94912	0.99663	0.99292	0.047102	0.0022931	0.0048104
		Sensitivity	0.49786	0.5406	0.52564	0.11325	0.15171	0.14103
✗	OI 2	Accuracy	0.91958	0.96587	0.97832	0.05523	0.0087228	0.0055179
		Specificity	0.92169	0.96918	0.9808	0.056776	0.0088467	0.0067348
		Sensitivity	0.50427	0.37607	0.50427	0.2735	0.068376	0.2735
	FS 5	Accuracy	0.94028	0.97339	0.98862	0.049141	0.011061	0.0014662
		Specificity	0.9443	0.9792	0.99423	0.049704	0.010533	0.0024401
		Sensitivity	0.53632	0.39316	0.42094	0.074786	0.16239	0.19017
	OI 5	Accuracy	0.95956	0.97694	0.97049	0.029203	7.2171e-05	0.0083567
		Specificity	0.96381	0.98177	0.97487	0.029521	0.00043532	0.010463
		Sensitivity	0.53632	0.50855	0.50427	0.074786	0.047009	0.2735
✗	FS 5	Accuracy	0.98163	0.99338	0.99365	0.013483	0.00089612	0.0025942
		Specificity	0.98393	0.99645	0.99639	0.014185	0.0017439	0.0032333
		Sensitivity	0.55342	0.46368	0.48718	0.1688	0.074786	0.17949

✗	OI 5	Accuracy	0.96938	0.97929	0.98348	0.02235	0.00040796	0.0007726
		Specificity	0.97196	0.98248	0.98618	0.022752	0.00023244	0.00022434
		Sensitivity	0.49786	0.42521	0.4765	0.11325	0.036325	0.24573
	FS 10	Accuracy	0.96661	0.99096	0.99164	0.026783	0.0011054	0.00042923
		Specificity	0.97059	0.9966	0.99866	0.028081	0.0020712	0.00065777
		Sensitivity	0.55342	0.42094	0.32051	0.1688	0.19017	0.012821
	OI 10	Accuracy	0.97472	0.98965	0.99063	0.01867	0.00088505	0.00077489
		Specificity	0.97916	0.99564	0.9973	0.018174	0.0023066	3.5592e-05
		Sensitivity	0.54701	0.42521	0.34829	0.008547	0.036325	0.040598
✗	FS 10	Accuracy	0.96517	0.99449	0.99434	0.029947	0.00050215	0.00077758
		Specificity	0.96718	0.99795	0.99745	0.030936	0.0009708	0.0014146
		Sensitivity	0.5812	0.38675	0.42094	0.19658	0.0021368	0.19017
✗	OI 10	Accuracy	0.98139	0.99449	0.99505	0.011094	0.00050215	6.176e-05
		Specificity	0.98405	0.99795	0.9987	0.011791	0.0009708	0.00021523
		Sensitivity	0.48718	0.38675	0.34829	0.17949	0.0021368	0.040598

Table F.10: Cross-validation results for randf, patient 39.

Method: Random Forest, Trees: 10, 30, 50								
KDE	Red		μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
		Accuracy	0.87302	0.89789	0.93672	0.056924	0.014864	0.020003
		Specificity	0.87493	0.90052	0.93882	0.059736	0.014144	0.021484
		Sensitivity	0.63675	0.65171	0.75641	0.25214	0.040598	0.089744
✗		Accuracy	0.98043	0.98664	0.992	0.012812	0.0005166	0.0019144
		Specificity	0.98464	0.99123	0.99632	0.013307	0.00055989	0.0023155
		Sensitivity	0.57479	0.53632	0.57479	0.036325	0.074786	0.036325
	FS 2	Accuracy	0.85579	0.88631	0.88907	0.074565	0.025112	0.032496
		Specificity	0.85747	0.8889	0.89169	0.076137	0.02623	0.033021
		Sensitivity	0.71795	0.65171	0.64103	0.051282	0.040598	0.025641
	OI 2	Accuracy	0.90236	0.93526	0.93814	0.0044518	0.0086226	0.0023741
		Specificity	0.90603	0.93867	0.94189	0.0059658	0.010675	0.0014408
		Sensitivity	0.51496	0.63462	0.58547	0.20726	0.13462	0.029915
✗	FS 2	Accuracy	0.93932	0.97083	0.96927	0.012679	0.021059	0.0016561
		Specificity	0.94209	0.97461	0.97369	0.013422	0.02128	0.0023552
		Sensitivity	0.67949	0.60256	0.54701	0.012821	0.064103	0.008547
✗	OI 2	Accuracy	0.96983	0.97874	0.97577	0.021382	0.011791	0.014765
		Specificity	0.97496	0.98462	0.98094	0.021622	0.011956	0.014271
		Sensitivity	0.47009	0.40385	0.4594	0.08547	0.096154	0.15171
	FS 5	Accuracy	0.848	0.88705	0.91567	0.035706	0.031136	0.020098
		Specificity	0.85188	0.88828	0.91787	0.03655	0.03095	0.020533
		Sensitivity	0.48077	0.76282	0.70726	0.019231	0.070513	0.014957
	OI 5	Accuracy	0.9293	0.94592	0.94462	0.013903	0.00607	0.0087289
		Specificity	0.93298	0.94841	0.94777	0.014315	0.0064127	0.0091061
		Sensitivity	0.57479	0.70726	0.64103	0.036325	0.014957	0.025641
✗	FS 5	Accuracy	0.95796	0.96231	0.98503	0.0062067	0.0065904	0.0061802
		Specificity	0.96228	0.96666	0.98861	0.0076085	0.0073285	0.0072799
		Sensitivity	0.55769	0.54701	0.65171	0.057692	0.008547	0.040598
✗	OI 5	Accuracy	0.96616	0.98171	0.97908	0.022343	0.0081478	0.012129
		Specificity	0.97096	0.9856	0.98293	0.025622	0.0089288	0.011595
		Sensitivity	0.55128	0.61325	0.59188	0.21795	0.0021368	0.13034
	FS 10	Accuracy	0.91859	0.95933	0.95464	0.041521	0.0088961	0.0081744
		Specificity	0.92051	0.96228	0.95721	0.041844	0.0082841	0.0078795
		Sensitivity	0.73504	0.65812	0.69658	0.042735	0.11966	0.081197
	OI 10	Accuracy	0.84547	0.87744	0.94092	0.048378	0.067792	0.0070158
		Specificity	0.84598	0.879	0.9437	0.049981	0.069666	0.0077016
		Sensitivity	0.82265	0.75641	0.67949	0.10043	0.089744	0.012821
✗	FS 10	Accuracy	0.98604	0.99196	0.992	0.0065258	0.0014282	0.0019144
		Specificity	0.98828	0.99527	0.99495	0.0076133	6.2286e-05	0.00094661
		Sensitivity	0.78419	0.69017	0.6859	0.061966	0.07906	0.14744
✗	OI 10	Accuracy	0.97734	0.98934	0.98865	0.0057585	0.0032211	0.0012079
		Specificity	0.98118	0.9943	0.99293	0.0071821	0.0029644	0.0015955
		Sensitivity	0.62393	0.49786	0.57479	0.068376	0.11325	0.036325

Table F.11: Cross-validation results for logistic regression, patient 39.

Method: Logistic Regression				
KDE	Red		μ_1	σ_1
		Accuracy	0.79471	0.1113
		Specificity	0.79886	0.11353
		Sensitivity	0.42521	0.036325
✗		Accuracy	0.87392	0.079555
		Specificity	0.87678	0.079273
		Sensitivity	0.3547	0.20085
	FS 2	Accuracy	0.84191	0.087776
		Specificity	0.84685	0.090181
		Sensitivity	0.39744	0.064103
	OI 2	Accuracy	0.91745	0.00073311
		Specificity	0.92167	0.0016701
		Sensitivity	0.51923	0.019231
✗	FS 2	Accuracy	0.91945	0.047267
		Specificity	0.92194	0.048879
		Sensitivity	0.51282	0.17949
✗	OI 2	Accuracy	0.97471	0.00047874
		Specificity	0.97734	0.00075169
		Sensitivity	0.49786	0.11325
	FS 5	Accuracy	0.83249	0.08299
		Specificity	0.83699	0.085661
		Sensitivity	0.4359	0.10256
	OI 5	Accuracy	0.91655	0.0097469
		Specificity	0.92248	0.011812
		Sensitivity	0.36966	0.09188
✗	FS 5	Accuracy	0.92085	0.049799
		Specificity	0.92412	0.050302
		Sensitivity	0.34829	0.040598
✗	OI 5	Accuracy	0.8599	0.0886
		Specificity	0.86144	0.089826
		Sensitivity	0.55342	0.1688
	FS 10	Accuracy	0.85732	0.0027235
		Specificity	0.86057	0.0039049
		Sensitivity	0.55769	0.057692
	OI 10	Accuracy	0.85443	0.016026
		Specificity	0.85728	0.015389
		Sensitivity	0.58547	0.029915
✗	FS 10	Accuracy	0.84004	0.047934
		Specificity	0.84304	0.048174
		Sensitivity	0.30983	0.07906
✗	OI 10	Accuracy	0.87063	0.070346
		Specificity	0.87297	0.070738
		Sensitivity	0.44231	0.057692

Table F.12: Cross-validation results for SVM, patient 39.

Method: SVM, Trees: 10, 30, 50										
KDE	Red		μ_1	μ_2	μ_3	μ_4	σ_1	σ_2	σ_3	σ_4
		Accuracy	0.98962	0.98962	0.94126	0.94126	0.0017891	0.0017891	0.0371	0.0371
		Specificity	1	1	0.94547	0.94547	0	0	0.039471	0.039471
		Sensitivity	0	0	0.56838	0.56838	0	0	0.12393	0.12393
✗		Accuracy	0.99434	0.99434	0.96365	0.96365	0.00077758	0.00077758	0.024179	0.024179
		Specificity	1	1	0.96582	0.96582	0	0	0.025891	0.025891
		Sensitivity	0	0	0.61752	0.61752	0	0	0.22863	0.22863
	FS 2	Accuracy	0.866	0.90232	0.95538	0.90391	0.027592	0.051029	0.014879	0.044028
		Specificity	0.86726	0.90716	0.96031	0.90875	0.02674	0.053826	0.013694	0.046079
		Sensitivity	0.75641	0.47436	0.49145	0.46368	0.089744	0.14103	0.047009	0.074786
	OI 2	Accuracy	0.9683	0.86488	0.91537	0.91903	0.0046532	0.050604	0.0066436	0.0076046
		Specificity	0.9734	0.86863	0.91955	0.92358	0.0060664	0.051968	0.0057878	0.0070856

		Sensitivity	0.49145	0.51923	0.51923	0.48077	0.047009	0.019231	0.019231	0.019231
✗	FS 2	Accuracy	0.96133	0.95228	0.97467	0.94541	0.029003	0.024458	0.0065285	0.027744
		Specificity	0.96646	0.95607	0.97693	0.94788	0.030302	0.025201	0.0068272	0.028347
		Sensitivity	0.076923	0.29274	0.5641	0.51923	0.076923	0.014957	0.10256	0.019231
✗	OI 2	Accuracy	0.94234	0.89892	0.97009	0.96774	0.018935	0.035287	0.00091819	0.00030745
		Specificity	0.94551	0.90074	0.97286	0.97015	0.019357	0.036547	0.0011987	0.00076503
		Sensitivity	0.36538	0.54274	0.48077	0.52564	0.13462	0.23504	0.019231	0.14103
	FS 5	Accuracy	0.98466	0.9846	0.94396	0.89211	0.0031679	0.0021803	0.0093899	0.064616
		Specificity	0.99467	0.99358	0.94916	0.89654	0.0053333	0.00042231	0.011826	0.067871
		Sensitivity	0.038462	0.13248	0.47436	0.51282	0.038462	0.021368	0.14103	0.17949
	OI 5	Accuracy	0.98431	0.974	0.93776	0.80062	0.0014928	0.0057205	0.0067233	0.074458
		Specificity	0.99464	0.98423	0.94319	0.80484	0.0033066	0.0075599	0.0054764	0.07383
		Sensitivity	0	0	0.42521	0.39744	0	0	0.036325	0.064103
✗	FS 5	Accuracy	0.99377	0.98785	0.96072	0.93685	0.00021367	0.0021321	0.016733	0.040606
		Specificity	0.99943	0.99292	0.9638	0.94	0.00056668	0.0031188	0.017826	0.041626
		Sensitivity	0	0.1047	0.4359	0.39744	0	0.049145	0.10256	0.064103
✗	OI 5	Accuracy	0.99083	0.98508	0.97108	0.94272	0.00022538	0.0027501	0.0052942	0.0042765
		Specificity	0.99647	0.9907	0.97478	0.94461	0.0010059	0.0035405	0.006132	0.0047676
		Sensitivity	0	0	0.3312	0.59188	0	0	0.053419	0.13034
	FS 10	Accuracy	0.98962	0.98896	0.94106	0.90441	0.0017891	0.0011282	0.019049	0.059074
		Specificity	1	0.99933	0.94522	0.9086	0	0.00066667	0.020555	0.0606
		Sensitivity	0	0	0.55769	0.51923	0	0	0.057692	0.019231
	OI 10	Accuracy	0.98962	0.98896	0.90802	0.86364	0.0017891	0.0011282	0.035805	0.02929
		Specificity	1	0.99933	0.91207	0.86789	0	0.00066667	0.035268	0.030109
		Sensitivity	0	0	0.51923	0.43162	0	0	0.019231	0.12393
✗	FS 10	Accuracy	0.99415	0.99415	0.96436	0.94425	0.00058961	0.00058961	0.010227	0.030339
		Specificity	0.99981	0.99981	0.96711	0.94633	0.00018889	0.00018889	0.010556	0.031334
		Sensitivity	0	0	0.48077	0.59615	0	0	0.019231	0.096154
✗	OI 10	Accuracy	0.99434	0.99396	0.91243	0.90409	0.00077758	0.00040164	0.055238	0.0068112
		Specificity	1	0.99962	0.9152	0.90649	0	0.00037779	0.055331	0.0069956
		Sensitivity	0	0	0.41453	0.47009	0	0	0.029915	0.08547

Table F.13: Post-processed results for KNN, patient 39.

Method: KNN, Neighbors: 2, 5, 10											
			Original			Veto 1			Veto 2		
KDE	Red		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
		TP	2	2	2	2	2	2	2	2	2
		FP	14	0	0	13	0	0	5	0	0
		FN	0	0	0	0	0	0	0	0	0
✗		TP	0	0	0	0	0	0	0	0	0
		FP	2	0	0	0	0	0	0	0	0
		FN	2	2	2	2	2	2	2	2	2
	FS 2	TP	2	2	2	2	2	2	2	1	2
		FP	25	54	9	21	36	8	10	26	4
		FN	0	0	0	0	0	0	0	1	0
	OI 2	TP	2	2	2	2	2	2	2	2	2
		FP	31	6	14	24	6	9	14	4	6
		FN	0	0	0	0	0	0	0	0	0
✗	FS 2	TP	0	0	0	0	0	0	0	0	0
		FP	11	4	3	2	2	0	2	2	0
		FN	2	2	2	2	2	2	2	2	2
✗	OI 2	TP	1	0	0	1	0	0	1	0	0
		FP	10	5	3	1	2	0	1	0	0
		FN	1	2	2	1	2	2	1	2	2
	FS 5	TP	2	2	2	2	2	2	2	1	2
		FP	33	13	2	23	9	2	11	1	0
		FN	0	0	0	0	0	0	0	1	0
	OI 5	TP	2	2	2	2	2	2	2	2	2
		FP	20	8	13	16	8	10	8	6	8
		FN	0	0	0	0	0	0	0	0	0
✗	FS 5	TP	0	0	0	0	0	0	0	0	0
		FP	4	2	1	0	1	0	0	0	0

		FN	2	2	2	2	2	2	2	2	2
		TP	0	0	0	0	0	0	0	0	0
✗	OI 5	FP	5	3	3	0	1	0	0	1	0
		FN	2	2	2	2	2	2	2	2	2
		TP	2	2	2	2	2	2	2	1	2
	FS 10	FP	14	1	0	9	0	0	5	0	0
		FN	0	0	0	0	0	0	0	1	0
		TP	2	2	2	2	2	2	2	2	2
	OI 10	FP	10	0	0	8	0	0	4	0	0
		FN	0	0	0	0	0	0	0	0	0
		TP	0	0	0	0	0	0	0	0	0
✗	FS 10	FP	8	1	2	1	0	0	0	0	0
		FN	2	2	2	2	2	2	2	2	2
		TP	0	0	0	0	0	0	0	0	0
✗	OI 10	FP	1	0	0	0	0	0	0	0	0
		FN	2	2	2	2	2	2	2	2	2

Table F.14: Post-processed results for random forest, patient 39.

Method: Random Forest, Trees: 10, 30, 50											
			Original			Veto 1			Veto 2		
KDE	Red		μ_1	μ_2	μ_3	μ_1	μ_2	μ_3	μ_1	μ_2	μ_3
		TP	2	2	2	2	2	2	2	2	2
		FP	51	48	28	28	37	25	20	26	15
		FN	0	0	0	0	0	0	0	0	0
		TP	2	2	2	2	2	2	2	2	2
✗		FP	3	4	1	3	1	1	2	1	0
		FN	0	0	0	0	0	0	0	0	0
		TP	2	2	2	2	2	2	2	2	2
	FS 2	FP	44	47	42	27	31	30	20	17	16
		FN	0	0	0	0	0	0	0	0	0
		TP	2	2	2	2	2	2	2	2	2
	OI 2	FP	47	32	33	31	26	26	19	11	13
		FN	0	0	0	0	0	0	0	0	0
		TP	2	2	2	2	2	2	2	2	2
✗	FS 2	FP	17	6	9	8	5	3	4	4	2
		FN	0	0	0	0	0	0	0	0	0
		TP	2	2	2	2	2	2	2	2	2
✗	OI 2	FP	5	7	5	4	6	3	2	3	1
		FN	0	0	0	0	0	0	0	0	0
		TP	2	2	2	2	2	2	2	2	2
	FS 5	FP	62	44	36	28	36	29	11	26	19
		FN	0	0	0	0	0	0	0	0	0
		TP	2	2	2	2	2	2	2	2	2
	OI 5	FP	31	25	28	25	20	22	16	11	12
		FN	0	0	0	0	0	0	0	0	0
		TP	2	2	2	2	2	2	2	2	2
✗	FS 5	FP	12	14	4	4	6	3	2	4	2
		FN	0	0	0	0	0	0	0	0	0
		TP	2	2	2	2	2	2	2	2	2
✗	OI 5	FP	7	5	4	6	2	2	4	1	1
		FN	0	0	0	0	0	0	0	0	0
		TP	2	2	2	2	2	2	2	2	2
	FS 10	FP	30	15	23	25	13	16	16	7	12
		FN	0	0	0	0	0	0	0	0	0
		TP	2	2	2	2	2	2	2	2	2
	OI 10	FP	56	44	28	46	31	24	26	17	18
		FN	0	0	0	0	0	0	0	0	0
		TP	2	2	2	2	2	2	2	2	2
✗	FS 10	FP	6	4	2	4	2	1	2	1	1
		FN	0	0	0	0	0	0	0	0	0
		TP	2	2	2	2	2	2	2	2	2
✗	OI 10	FP	5	3	4	1	2	3	1	1	1
		FN	0	0	0	0	0	0	0	0	0

Table F.15: Post-processed results for logistic regression, patient 39.

Method: Logistic Regression					
KDE	Red		Original	Veto 1	Veto 2
		TP	2	2	2
		FP	73	31	16
		FN	0	0	0
✗		TP	0	0	0
		FP	11	5	2
		FN	2	2	2
	FS 2	TP	2	2	1
		FP	52	34	19
		FN	0	0	1
	OI 2	TP	2	2	2
		FP	33	27	21
		FN	0	0	0
✗	FS 2	TP	0	0	0
		FP	10	5	4
		FN	2	2	2
✗	OI 2	TP	0	0	0
		FP	8	4	2
		FN	2	2	2
	FS 5	TP	2	2	2
		FP	59	37	18
		FN	0	0	0
	OI 5	TP	2	2	2
		FP	37	24	17
		FN	0	0	0
✗	FS 5	TP	1	1	1
		FP	11	3	2
		FN	1	1	1
✗	OI 5	TP	1	1	0
		FP	17	2	1
		FN	1	1	2
	FS 10	TP	2	2	2
		FP	63	41	18
		FN	0	0	0
	OI 10	TP	2	2	2
		FP	63	31	17
		FN	0	0	0
✗	FS 10	TP	2	2	1
		FP	15	5	3
		FN	0	0	1
✗	OI 10	TP	1	1	0
		FP	17	3	2
		FN	1	1	2

Table F.16: Post-processed results SVM, patient 39.

Method: SVM, kernel: rbf, linear, BoxConst: 1, 100														
			Original				Veto 1				Veto 2			
KDE	Red		μ_1	μ_2	μ_3	μ_4	μ_1	μ_2	μ_3	μ_4	μ_1	μ_2	μ_3	μ_4
		TP	0	0	2	2	0	0	2	2	0	0	2	2
		FP	0	0	22	22	0	0	17	17	0	0	11	11
		FN	2	2	0	0	2	2	0	0	2	2	0	0
✗		TP	0	0	1	1	0	0	1	1	0	0	1	1
		FP	0	0	6	6	0	0	5	5	0	0	3	3
		FN	2	2	1	1	2	2	1	1	2	2	1	1
	FS 2	TP	2	2	2	2	2	2	2	2	2	2	2	1
		FP	45	38	24	41	36	27	16	26	27	12	11	20
		FN	0	0	0	0	0	0	0	0	0	0	0	1

	OI 2	TP	2	2	2	2	2	2	2	2	2	2	2	2
		FP	12	55	37	42	11	39	29	31	6	17	24	20
		FN	0	0	0	0	0	0	0	0	0	0	0	0
✗	FS 2	TP	0	0	0	0	0	0	0	0	0	0	0	0
		FP	4	4	7	10	3	2	2	5	1	2	0	4
		FN	2	2	2	2	2	2	2	2	2	2	2	2
✗	OI 2	TP	1	1	0	0	0	1	0	0	1	0	0	0
		FP	7	8	9	8	1	2	3	4	1	1	3	3
		FN	1	1	2	2	2	1	2	2	1	2	2	2
	FS 5	TP	0	2	2	2	0	2	2	2	0	1	2	2
		FP	2	0	28	42	0	0	21	28	0	0	10	15
		FN	2	0	0	0	2	0	0	0	2	1	0	0
	OI 5	TP	0	0	2	2	0	0	2	2	0	0	2	2
		FP	1	5	28	65	1	5	25	37	0	0	14	27
		FN	2	2	0	0	2	2	0	0	2	2	0	0
✗	FS 5	TP	0	0	0	0	0	0	0	0	0	0	0	0
		FP	0	2	5	7	0	1	4	5	0	1	3	4
		FN	2	2	2	2	2	2	2	2	2	2	2	2
✗	OI 5	TP	0	0	0	1	0	0	0	1	0	0	0	1
		FP	0	0	7	13	0	0	4	5	0	0	2	4
		FN	2	2	2	1	2	2	2	1	2	2	2	1
	FS 10	TP	0	0	2	2	0	0	2	2	0	0	2	2
		FP	0	0	30	42	0	0	23	21	0	0	13	10
		FN	2	2	0	0	2	2	0	0	2	2	0	0
	OI 10	TP	0	0	2	2	0	0	2	2	0	0	2	2
		FP	0	0	46	56	0	0	23	29	0	0	13	17
		FN	2	2	0	0	2	2	0	0	2	2	0	0
✗	FS 10	TP	0	0	0	0	0	0	0	0	0	0	0	0
		FP	0	0	7	9	0	0	4	5	0	0	1	5
		FN	2	2	2	2	2	2	2	2	2	2	2	2
✗	OI 10	TP	0	0	1	1	0	0	1	1	0	0	0	0
		FP	0	0	15	15	0	0	2	5	0	0	2	2
		FN	2	2	1	1	2	2	1	1	2	2	2	2

G

Appendix 7: GTCS Seizure data

Table G.1: GTCS statistics describing seizure onsets, duration, pre-processed duration and standard deviation.

Patient	Start	Stop	Seizure time	Max STD
7	2008-09-12 01:39:19	2008-09-12 01:40:24	01:05	0.9105
7	2008-09-12 02:27:33	2008-09-12 02:28:38	01:05	0.7794
7	2008-09-12 03:18:07	2008-09-12 03:18:55	00:48	0.9730
7	2008-09-12 04:10:15	2008-09-12 04:11:30	01:15	0.9850
7	2008-09-12 04:57:25	2008-09-12 04:58:30	01:05	0.9499
7	2008-09-12 06:36:23	2008-09-12 06:37:23	01:00	1.0004
7	2008-09-12 08:02:13	2008-09-12 08:03:10	00:57	0.7671
7	2008-09-12 09:10:30	2008-09-12 09:11:39	01:09	0.9995
7	2008-09-12 10:54:45	2008-09-12 10:55:55	01:10	0.8906
7	2008-09-12 16:08:43	2008-09-12 16:09:47	01:04	0.9057
7	2008-09-12 18:18:49	2008-09-12 18:19:39	00:50	0.6503
22	2009-10-09 00:33:35	2009-10-09 00:34:37	01:02	0.6873
22	2009-10-09 01:20:50	2009-10-09 01:21:53	01:03	0.6453
22	2009-10-09 03:11:45	2009-10-09 03:12:51	01:06	0.9897
22	2009-10-09 03:54:00	2009-10-09 03:55:06	01:06	0.9953
22	2009-10-09 04:22:03	2009-10-09 04:23:16	01:13	0.6233
22	2009-10-09 04:47:37	2009-10-09 04:48:52	01:15	0.6622
22	2009-10-09 05:08:20	2009-10-09 05:09:34	01:14	0.6438
22	2009-10-09 05:50:52	2009-10-09 05:52:03	01:11	0.6297
48	2012-05-03 18:59:02	2012-05-03 19:00:13	01:11	0.8280
48	2012-05-03 21:41:23	2012-05-03 21:43:13	01:50	0.7839
55	2014-12-11 02:51:06	2014-12-11 02:52:17	01:11	1.4824
55	2014-12-11 09:59:26	2014-12-11 10:00:15	00:49	1.8422