VPH2

Virtual Pathological Heart of the Virtual Physiological Human

Grant Agreement Number 224635

– Deliverable –

D3.4 – Application of data mining methodologies

Document Information

Document Name:	**D3_4_CTI_WP3_067-09-2010_Final**
Revision:	1.1
Revision Date:	06.09.2010
Author:	CTI
Security:	Public

Document History

Revision	Date	Modification	Author
Version 0.1	01/06/2010	First Draft of ToC to be approved by involved partners	CTI
Version 0.2	07/06/210	ToC to be finalised and approved by the Board	CTI, all involved partners
Version 0.3	30/06/2010	First contributions	All involved partners
Version 0.4	10/06/2010	Complementary contributions	All involved partners
Version 0.5	15/07/2010	Modifications	All involved partners
Version 0.6	15/07/2010-15/08/2010	Delay due to the availability of the Niguarda data	CTI, Q&R, CNR
Version 0.7	25/08/2010	Deliverable to be approved by Quality Mechanism	All
Version 1.0	31/08/2010	Refinement	CTI
Version 1.1	06/09/2010	Final Draft to be released	CTI

Table of contents

Abbreviations

The following table presents the main terms and acronyms used in this document.

CAD	Coronary Artery Disease
CHF	Chronic Heart Failure
AMI	Acute Myocardial Infarction
FAT	Functional Assessment Tool
LVD	Left Ventricle Dysfunction
MRI	Magnetic Resonance Image
DSS	Decision Support System
DB	Database
NYHA	New York Heart Association
PUFA	Polyunsaturated Fatty Acids
ACE	Angiotensin-Converting Enzyme
HDL	High-density lipoprotein
SGOT	Serum Glutamic Oxaloacetic Transaminase
SGPT	Serum Glutamic Pyruvic Transaminase
LV	Left Ventricle
HF	Heart Failure
LVEF	Left Ventricle Ejection Fraction
IHD	Ischeamic Heart Disease
CRT	Cardiac Resynchronization Therapy
STEMI	ST-Segment Elevation Myocardial Infarction
NSTEMI	Non-ST-Segment Elevation Myocardial Infarction
COPD	Chronic Obstructive Pulmonary Disease
RR	Relative Risk
CI	Confidence Interval
CIHD	Chronic Ischeamic Heart Disease
cIHF	Chronic Ischeamic Heart Failure
ARB	Angiotensin Receptor Blockers
ROC	Receiver Operating Characteristic
CABG	Coronary Artery Bypass Graft Surgery
XML	Extensible Markup Language
ADL	Archetype Definition Language
DBMS	Database Management Schema
GUI	Graphic User Interface
EHR	Electronic Health Record
DICOM	Digital Imaging and Communications in Medicine
AJAX	Asynchronous JavaScript and XML
FPT	Functional Predictive Tool
CVD	Cardio-vascular Disease

1. Introduction

Deliverable 3.4 is based on project task T3.2 – 'Build a framework system on Data mining and available for heterogeneous dataset' and it is the report describing the work performed during this task.

In the activities described below various VPH2 partners coming from multiple disciplines (data mining engineers, software engineers, clinicians, data base experts etc) were involved:

- CTI, Intercon and Q&R that worked mainly for the application of data mining methodologies in the data set and for developing a decision support software addressed to cardiologists.
- Niguarda and CNR that provided the clinical feedback, interpreting the results, asking for certain output from the available data set and providing valuable input for decreasing the variables and improving the accuracy of the results.
- WWU that provided feedback for any results containing data from the genetic study

In this deliverable the work done with Mario Negri dataset coming from GISSI Prevenzione study (available from the beginning of 2010) as well as the first results from the work with Niguarda dataset (available from 15 July 2010) are described. The GISSI dataset is anyway the most complete and the one that is coupled with the results from the genetic analysis conducted by WWU, it also sets the framework around which the decision support module of VPH2 project will be built. Actually this deliverable can be considered as the first version of the "Application of data mining techniques". A second and final version of this deliverable will be released at the end of the work package 3, i.e. month 30 of the project, i.e. December 2010. At that time all the data mining work with the various datasets will be thoroughly described and the two versions will constitute the description of the data mining activities during VPH2.

The GISSI study data set is described elsewhere (D2.4, D3.3) and there is no point in providing the same information again in this deliverable. On the other hand the Niguarda data set is described in chapter 5.1.b.

It should also be mentioned that this deliverable was delayed for a month in order to include some first, yet indicative results of the (ongoing) work done with Niguarda dataset. Niguarda retrospective data collection was actually a voluntary work (no commitment for IFC in the DoW) carried on at IFC with the administrative personnel of the hospital (no effort claimed). The number of patients was much higher than that supposed at the beginning, although in only 50% of the population EF and volumes have been collected. The population consists of patients suffering from chronic CAD with CHF and AMI with CHF. These patients can be matched with an equivalent population of chronic CAD or AMI without CHF. The total number of cases that were extracted was 2097. The data extraction for all took time, and the effort was taken by a voluntary work of the IFC researchers, as already explained. These data became available within the middle of July. The data mining work, always in close cooperation with and consulting from the involved CNR clinicians started immediately and the initial results are presented in the following (chapter 5.5).

The reasons why the other two data sets are not available is different in each case. Specifically:

- The 100 cases of patients for which the cardiac MRI examination was also extracted are available. But the FAT software which is necessary for the extraction of the features from this MRI examination will be released in its final, stable version by the end of June which means that the time needed for the processing of the MR images and then for the application of data mining in the resulting, enriched data set wasn’t enough. Consequently, this work will take place during July and August 2010 and the relevant activities will be described in the final release of this deliverable in December 2010.
- CRT data: the agreement with the clinical partners (Pavia, Rozzano, Niguarda) providing the blood samples needed for the genotyping study phase II, was to utilize their clinical data when the genetic data will be available, in order this deal to be profitable for all involved actors. i.e. samples providers and VPH2 partners. The clinical data are available at IFC and will take some more time to combine the 2 dataset (genetic and clinical). To sum up, once the genotyping analysis phase II is completed and the results are coupled with the existing clinical data the data mining work can start.

2. Project Overview

The VPH2 project aims to develop a patient-specific computational model and simulation of the human heart to assist cardiologists and cardiac surgeons in defining the severity and extent of disease in patients with Left Ventricular Dysfunction (LVD), with or without mitral regurgitation. Associated specific computational methods will allow clinical decision making and planning of the optimal treatment for left ventricle-valve repair.

The associated technological aim of the project is to deliver the most advanced software application framework for the development of computer-aided medicine in cardiology and cardiac surgery available in the world, going beyond the state of the art of available models.

This goal will be achieved by integrating some of the leading Open Source software in the area of computer-aided medicine and of computational bioengineering. This framework will be used by VPH2 to realise its objectives, but also by any other future project (academic or industrial) aiming to improve or extend VPH2 objectives.

3. The Role of this task in the VPH2 project

The related tasks are:

- Task 2.2 'Specification of the user requirements from a technical perspective', since the approach adopted has taken into account the users' needs, e.g. concerning the transparency of the methodologies and the user specific modification of the module.
- Task 3.1 – 'Analysis of the existing clinical databases', since the data sets available for the project were identified and thoroughly described so as to explore the possibilities of data mining and knowledge extraction in these data sets.
- Task 6.1 'Integration of the VPH2 framework using Spiral Approach'; the data mining work will be used for the development of the decision support module provided by VPH2 platform. This module will be integrated with Core DB so as to retrieve patients' data from it and store data back to it. Moreover it will be integrated with the experts' interfaces, i.e. the interfaces that will expose the functionality of this decision support module
- Task 6.2 'Development of Front-End application'. This task includes the implementation of experts' interfaces through which the extracted knowledge and the decision support functionalities will be exposed to the end users, i.e. to the clinicians.

4. The Role of this deliverable in the VPH2 project

The role of this deliverable is the detailed description of all activities concerning the whole data mining/ knowledge extraction process, i.e. data availability, data management etc. Within this document technologies are specified and functions are thoroughly described. Any information necessary for future upgrades of these individual VPH2 modules is provided by this document.

The complete picture will be drawn in the second version of this deliverable at the completion of WP3 at the end of December 2010.

5. Data Mining

5.1 Rationale and Description of datasets

a. GISSI Study

Development of heart failure (HF) after acute myocardial infarction (AMI) is common: the incidence of in-hospital heart failure after the acute event varies between 18% in databases of clinical trials up to 37% in community studies [1]. Remodelling of the left ventricle (LV) as determined by echocardiography in the absence of overt HF is likewise very common, even in the current primary angioplasty era: up to one-third of patients with successful revascularization and sustained patency of the infarct-related artery, presented LV remodelling 6 months after acute MI [2].

Early HF after AMI is related to extensive myocardial damage, and thus to the severity of myocardial infarction. In contrast, late HF during follow-up correlates to the extent and the severity of the LV remodelling process. In the CARE study [3] among stable AMI survivors with no previous history of HF, 6.3% had a subsequent HF admission within 5 years; the cumulative incidence of HF increased by 1.3% per year.

The strongest independent predictors of HF development are age, gender, diabetes and LV dysfunction after AMI: for each 1% decrease in baseline LV ejection fraction (LVEF), the risk of HF occurrence increases by 4%.

The extremely high incidence of new onset HF after AMI (40% at a median follow-up of 6 years) in a well-characterized community cohort [4], and its impressive fatality rate with a median survival of 4 years after diagnosis, underscore the clinical relevance of post AMI remodelling and its burden for the National Health Systems.

This work is very promising since, to the best of our knowledge, no previous study addressed the issue of data mining based knowledge extraction in a population with post-MI development of myocardial remodelling. Homogenous and rigorously collected datasets, such as those available from randomized clinical trials, are obviously best suited for this analysis. To achieve this goal a selected patient series, enrolled in the nineties in a randomized controlled trial [5] on the efficacy of unsaturated fatty acids in preventing mortality after MI (GISSI Prevenzione) is analyzed.

In a second stage, by blending baseline anonymous individual patient records to genetic variation information, cohort data has entered into the VPH2 modelling system, as simulation of source data to guide decision-making in the Virtual Pathological Heart with post-ischemic LVD. The study protocol has been approved by the Biobank Committee of Instituto di Ricerche Farmacologiche Mario Negri on November 27th, 2008.

Study Population

The study included patients enrolled in the GISSI Prevenzione trial, according to the following eligibility criteria as depicted in the following

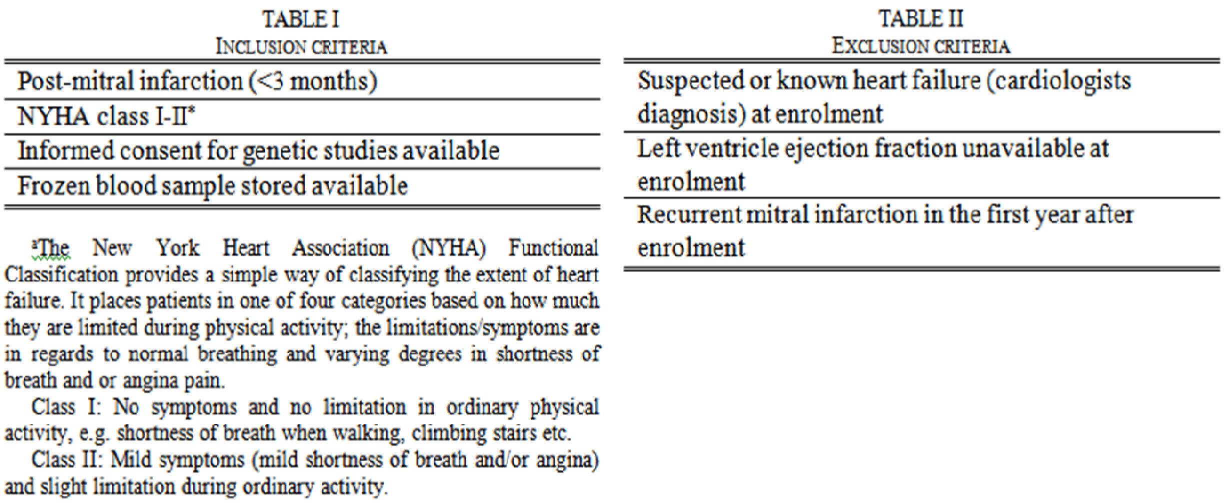


TABLE I
INCLUSION CRITERIA

Post-mitral infarction (<3 months)
NYHA class I-II*
Informed consent for genetic studies available
Frozen blood sample stored available

[a]The New York Heart Association (NYHA) Functional Classification provides a simple way of classifying the extent of heart failure. It places patients in one of four categories based on how much they are limited during physical activity; the limitations/symptoms are in regards to normal breathing and varying degrees in shortness of breath and or angina pain.

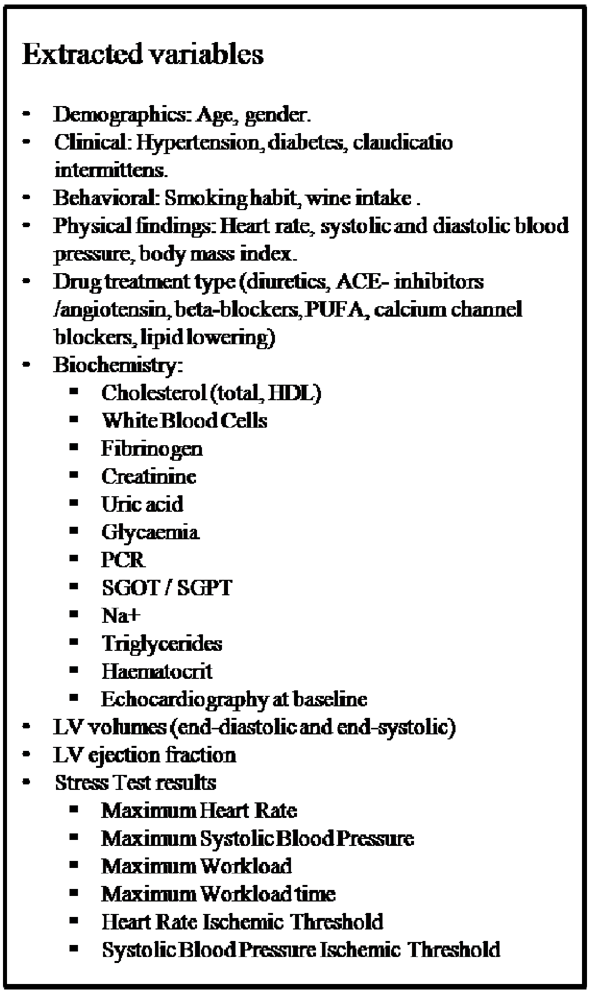
Class I: No symptoms and no limitation in ordinary physical activity, e.g. shortness of breath when walking, climbing stairs etc.

Class II: Mild symptoms (mild shortness of breath and/or angina) and slight limitation during ordinary activity.

TABLE II
EXCLUSION CRITERIA

Suspected or known heart failure (cardiologists diagnosis) at enrolment
Left ventricle ejection fraction unavailable at enrolment
Recurrent mitral infarction in the first year after enrolment

Patients meeting the above criteria were retrospectively identified from the GISSI Prevenzione database and the variables that were used after the cleansing of the dataset for those patients are depicted in Figure 1 below.



Extracted variables

- Demographics: Age, gender.
- Clinical: Hypertension, diabetes, claudicatio intermittens.
- Behavioral: Smoking habit, wine intake .
- Physical findings: Heart rate, systolic and diastolic blood pressure, body mass index.
- Drug treatment type (diuretics, ACE- inhibitors /angiotensin, beta-blockers, PUFA, calcium channel blockers, lipid lowering)
- Biochemistry:
  - Cholesterol (total, HDL)
  - White Blood Cells
  - Fibrinogen
  - Creatinine
  - Uric acid
  - Glycaemia
  - PCR
  - SGOT / SGPT
  - Na+
  - Triglycerides
  - Haematocrit
  - Echocardiography at baseline
- LV volumes (end-diastolic and end-systolic)
- LV ejection fraction
- Stress Test results
  - Maximum Heart Rate
  - Maximum Systolic Blood Pressure
  - Maximum Workload
  - Maximum Workload time
  - Heart Rate Ischemic Threshold
  - Systolic Blood Pressure Ischemic Threshold

Figure 1: Patients' data extracted from GISSI dataset

Genetics

Patients genomic DNA was extracted from mononuclear blood cells and screened for genetic variations. A high throughput genotyping approach using TaqMan assays in a 384-well ABI 7900HT Sequence Detection System (Applied Biosystems, Foster City, CA) was applied [6]. Primer and probe sequences are available upon request. Candidate genes and linked variants screened in this study were the result of a multilayer process considering most recent consolidated findings in clinical and molecular genetics of cardio-vascular dysfunction (CVD) with special respect to their reproducibility [7]. Since CVD is a multi-factorial trait being strongly genetically determined with a complex pathophysiology, the included genes represent five of the known major biological systems involved:

1) Adrenergic receptor system

2) Renin-Angiotensin-Aldosterone system

3) Endothelin system

4) Extracellular matrix enzymes

5) Inflammatory cytokines and cell adhesion molecules

Relative frequencies of genotypes and alleles have been compared by chi-square test (exact test when appropriate) in index cases (LVEF at baseline ≤ 0.40 or HF during follow-up) vs controls. P-values < 0.05 have been considered statistically significant without any multiple comparisons adjustment. The software gPLINK [8] has been used.

The most interesting results from data mining in the dataset enriched with the results of the genetic analysis are depicted in Chapter 5.4.

Data Preparation /Cleansing

Two case-control studies were conducted:

1) Cases were patients with baseline LVEF ≤ 0.40 (strong indication of heart remodelling), selected according to the above inclusion/exclusion criteria. Controls are patients with baseline LVEF > 0.40 (most probably not presenting heart remodelling) who did not develop late-onset HF during the whole follow-up period, matched in a 1:1 ratio to cases for age and gender. The total number of available samples was 1228.

2) Cases were patients who developed late-onset HF and were hospitalized for a clinical diagnosis of HF. Controls are patients who did not develop HF during the whole follow-up period, matched in a 1:1 ratio to cases for age and gender. The total number of samples was 202.

The data cleansing approach is described in the following:

1. Categorization of all patients in two main categories (those that developed late onset heart failure against those that did not develop it).
2. Features having more than 25% missing values are removed, such as stress test results.
3. Features that show no variation in their values are removed. Also features that are used to compute another feature are removed, such as LV end systolic, end diastolic volume, which are used to compute ejection fraction.
4. From the genetic data variants rs4291, rs5443 and rs4646994 were used (the p-values for these variants were 0.036, 0.0487 and 0.033 respectively). More details can be found in D4.2 where the association between these variants and late onset HF is thoroughly explained.
5. Due to imbalance of the dataset the SMOTE [9] algorithm was applied using ten nearest neighbours to create balanced datasets.

All common data mining methodologies (shortly presented in section 5.3 below) were applied in the cleansed data set, with the aim of classifying the patients according to the first case-control study: individuals where remodelling was observed against those that did not develop this feature.

Since, from a clinical point of view, this classification was not the most important one for this study, it was not further developed and was used much more as an exercise of collaboration between engineers and clinicians for the second study.

The second case control study aimed at classifying the patients in those that developed late onset heart failure against those that did not develop it. Since the initial results were not encouraging in terms of accuracy on one side and clinical interpretation on the other, the involved clinicians were asked to define more controlled datasets possibly increasing the chances of extracting any new knowledge.

Following clinicians' suggestions, 7 classifiers were built:

1. Diabetes, Ejection Fraction, AMI; these are the three more important variables that evidently affect the development of late onset heart failure
2. Diabetes, Ejection Fraction, AMI, Biochemical; in order to assess what lab data (i.e. cholesterol, white blood cells, fibrinogen, creatinine, uric acid) in general (and which one in particular) add in the predictive accuracy for late onset heart failure
3. Diabetes, Ejection Fraction, AMI, Genetics; in order to assess in general the genetic analysis results effect on the prediction of late onset heart failure
4. Diabetes, Ejection Fraction, AMI, PUFA; in order to assess what PUFA treatment adds in predictive accuracy of late on set heart failure
5. Genetics when Ejection Fraction > 40; in order to assess if genetic polymorphisms add predictive accuracy in healthy people
6. Genetics when Non Diabetic; in order to assess whether genetic polymorphisms add predictive accuracy only in non diabetic patients
7. Genetics when gender is female and age < 60 or gender is male and gender < 55; in order to assess if genetic polymorphisms add predictive accuracy in younger patients.

Multivariate analysis was performed using the Cox proportional risk model with the main aim being the determination of the indicators of LOHF in a large population of low risk survivors of AMI and to determine the prognosis of patients with this complication once diagnosed. The secondary aim was to determine the predictors of the composite event of death/LOHF. This work was part of the GISSI study (VPH2 had anyway access only in the 1228 cases with the DNA and not the whole dataset) and more details can be found in [10].

The most important classifiers are clinically interpreted (rule by rule in the most interesting findings) in section 5.5 – Clinicians feedback.

b. Niguarda dataset

Besides the GISSI dataset, including patients randomized to treatment with polyunsaturated fatty acids or placebo after an AMI in the early nineties, data mining was performed in a dataset of retrospectively enrolled real world ischemic heart disease patients.

The rationale for the choice of this population was to:

- derive real-world contemporary data (see below), obtained at the same location where prospective enrolment is ongoing, to populate the platform with data mining results
- have another AMI population to be compared with/to integrate GISSI (AMI trial patients) data
- have a population with chronic IHD and/or chronic ischemic heart failure who had undergone interventional (angioplasty and/or stenting or CRT or coronary surgery or valvular procedures) to be matched with prospectively enrolled patients
- have an hard end-point, i.e. vital status, as outcome during long-term (>1 year) follow-up

<u>Study population</u>

This dataset includes all patients admitted to Niguarda 2005 to 2008 with a clinical diagnosis of acute (AMI) or chronic IHD, for acute events or planned procedures, discharged alive, with the exclusion of patients who developed IHD in a transplanted heart.

Clinical data were retrieved from current hospital databases and manually checked as needed. No blood samples are available for retrospective genotyping in this population. Outcome data (vital status at an average follow-up of 3.5 years) was derived from census.

Patient records with no more than 25% missing values for demographic, clinical, echocardiography, laboratory and drug therapy data were considered for analysis.

Two separate datasets based on clinical diagnosis were examined:

- patients admitted for AMI (974 cases)
- patients admitted for chronic ischemic heart disease or chronic ischemic heart failure (404 cases)

The targeted outcome in both cases is the survival of the patients.

Retrospective AMI data set

Trends from published epidemiological studies and clinical trials show that in past 20 years, pharmacological and interventional therapies in AMI have changed substantially with associated decreases in-hospital complications and early case fatality in all infarction types [11-14]. These positive changes in outcome occurred despite a higher risk profile of patients presenting with AMI, who are in general older and have more frequently history of diabetes, hypertension, current smoking, heart failure, prior revascularization, stroke, and hyperlipidemia. Improvements in outcome have been associated to early reperfusion strategies and are apparent even in the older population strata [15]. Consistent prognostic predictors in the literature include age, gender, type of AMI (STEMI vs NSTEMI), comorbidities such as diabetes, anaemia, renal dysfunction, associated non-coronary vascular disease, incident heart failure, left ventricular ejection fraction, diabetes, atrial fibrillation, statin treatment [16-19].

The AMI retrospective data set includes 974 patients median age 67 years, 39% women, death rate 12.6%. Of these 48% were current or previous smokers, 57% had a history of hypertension, 21% of diabetes, 39% of dyslipidemia, 18% of chronic kidney dysfunction, 6% of atrial fibrillation, 10% of peripheral or cerebrovascular disease. AMI type was STEMI in 76%, 73% of patients underwent a primary percutaneous coronary intervention with stent implantation (in more than 1 vessel in 31% of these), while 7% overall underwent coronary artery bypass grafting during the index admission. Median LVEF was 55%. LV systolic dysfunction (LVEF <45%) was present in 19% and LV dilation in 26% of patients with LV dimensions recorded. Clinical evidence of HF was found in 20%. At discharge statins were prescribed to 79% of patient. The clinical profile of the population is therefore consistent with published data.

The following variables, shown to be predictive of outcome in the literature, were analysed

Table 1: Variables from AMI dataset (Niguarda)

Demographics	Clinical	Laboratory	Treatment
Age	Hypertension	Blood Glucose (Serum)	ACE - Inhibitors
Sex	Diabetes	Creatinine	Angiotensin-Receptor Blockers
Body Mass Index	Dyslipidemia	Haematocrit	Beta Blockers
Smoking Habit	Chronic kidney dysfunction	Haemoglobin (blood)	Calcium Channel Blockers
	Atrial fibrillation (chronic, transient)	PCR	ASA (AcetylSalicylic Acid)
	Pre-Existing Vascular Disease	Serum Total Cholesterol	Double Antiplatelet
	AMI Type	Triglycerides	Clopidogrel
	AMI Site	Troponin - T	Aldosterone Antagonists
	N vessels	Urea	Hypoglycaemic agents
	STENT	Uric Acid	Insulin
	STENT	Ves 1h	Statins

	Echocardiographic LV dilation	Leukocytes	Loop Diuretics
	LV Ejection Fraction		PUFA (ω-3)
	HF signs or symtpoms		

To confirm the consistency of this retrospectively enrolled series with AMI populations described in the literature, CNR analysed the predictive value of recorded variables by classical Cox proportional hazards models. The variables described in Table below significant by invariable analysis where consecutively entered in multi-variables models in blocks of

1. Demographics

2. Clinical

3. Laboratory

4. Treatment

Independent predictors of outcome (all-cause mortality) identified through a forward selection procedure were

Variable	RR	95% CI
Age	1.067	1.047 -1.086
Chronic kidney dysfunction	1.556	1.042 -2.326
COPD	1.939	1.191 -3.158
Peripheral Vascular Disease	1.830	0.910 -3.682
Cerebrovascular Disease	1.694	0.899 -3.194
Both Peripheral and Cerebrovascular Disease	4.796	2.403 -9.573
LV Ejection Fraction	0.973	0.957 -0.989
Calcium Channel Blockers	1.951	1.225 -3.108
Insulin	2.120	1.219 -3.685
Statins	0.475	0.326 -0.691

These findings are consistent with the published literature, whereby older age; comorbid conditions (with insulin treatment to be considered a proxy for complicated diabetes) are associated to a worse outcome and better systolic function and statin treatment to a better prognosis. Our results suggest that these retrospective individual patient data may reflect larger series, are representative of contemporary patients admitted to hospitals for AMI and are appropriate to populate the platform. The application of data mining methods may therefore derive rules that improve clinician decision-making.

Retrospective chronic ischemic heart disease (cIHD) and chronic ischemic heart failure (cIHF) data set

Ischemic heart disease is nowadays the commonest cause of heart failure in western countries. Consistent prognostic predictors in the literature include age, comorbidities such as diabetes, atrial fibrillation, anaemia, renal dysfunction, left ventricular ejection fraction and volumes, drug treatment such as ACE- or ARB-inhibitors, beta-blockers, and loop diuretics.

This dataset includes overall 404 patients of whom 172 had chronic ischemic heart disease (cIHD) and 232 chronic ischemic heart failure (cIHF); median age was 66 years, 16% were women, the overall death rate was 17%. 38% were current or previous smokers, 53% had a history of hypertension, 32% of diabetes, 42% of dyslipidemia, 22% of chronic kidney dysfunction, 13% of atrial fibrillation, 14% of peripheral or cerebrovascular disease. 16% of patients underwent a percutaneous coronary intervention with stent implantation (in more than 1 vessel in 44% of these), while 30% overall underwent coronary artery bypass grafting during the index admission. Median LVEF was 40%. LV   dilation was present in 60% of patients with LV dimensions recorded at discharge statins were prescribed to 69% of patients.

When compared to cIHD patients, cIHF subjects had a 4-fold higher mortality rate, severely depressed ventricular function and dilation, higher proportion of diabetics, and lower proportion of dyslipidemia, statin prescription

Table 2: Variables from chronic dataset (Niguarda)

Demographics	Clinical	Laboratory	Treatment
Age	Hypertension	Glucose	ACE - Inhibitors
Sex	Diabetes	Creatinine	Angiotensin-Receptor Blockers
Body Mass Index	Dyslipidemia	Haematocrit	Beta Blockers
Smoking Habit	Chronic kidney dysfunction	K	Calcium Channel Blockers
	Atrial fibrillation (chronic, transient)	NA	Aldosterone Antagonists
	Pre-Existing Vascular Disease	Total Bilirubine	Statins
	Previous STENT	Urea	Loop Diuretics
	N vessels	Uric Acid	Loop diuretics dose
	LV end-Diastolic Volume		CABG index admission
	LV end-Systolic Volume		Number bypass
	LV Ejection Fraction		Biventricular pacing
	HF signs or symtpoms		Implantable Cardioverter defibrillator

To confirm the consistency of this retrospectively enrolled series with chronic ischemic populations described in the literature, CNR analysed the predictive value of recorded variables by classical Cox proportional hazards models. The variables described in Table below significant by invariable analysis where consecutively entered in multivariable models in blocks of

1. Demographics

2. Clinical

3. Laboratory

4. Treatment

Independent predictors of outcome (all-cause mortality) identified through a forward selection procedure were

Variable	RR	95%	CI
Age	1.051	1.026-	1.077
Diabetes	2.086	1.288-	3.377
CABG index admission	0.355	0.159-	0.794
BetaBlockers	0.415	0.250-	0.688
Loop diuretics dose	1.006	1.003-	1.008
cl_groups cIHF vs cIHD	3.564	1.752-	7.249

These findings are again consistent with the published literature, whereby older age, diabetes and higher doses of loop diuretics (as proxy for persistent or worsening congestion) are associated to a worse outcome and beta-blocker treatment and surgical revascularization for the relief of ischemia to a better prognosis. Our results suggest that these retrospective individual patient data may reflect larger series, are representative of contemporary patients admitted to hospitals for cIHD/cIHF and are appropriate to populate the platform.

Data Preparation /Cleansing

Two studies were conducted, since the dataset was actually split in two main subsets according to the disease the patients suffered from:

1) Patients admitted for AMI (974 cases) with the outcome being the survival of these patients.

2) Patients admitted for chronic ischemic heart disease or chronic ischemic heart failure (404 cases) with the outcome being the survival of these patients.

As mentioned above the target variable was the survival of the patients. For that specific reason the data cleansing approach was the following:

1) Categorization of all patients in two main categories (those that are still alive against those that died).

2) Features selection according to clinicians feedback (in order to save time and work with a more concise and straightforward dataset). More details are provided in section 5.4 Results were the variables used for each classifier are explained.

3) Due to imbalance of the dataset the SMOTE [9] algorithm was applied using ten nearest neighbours to create balanced datasets.

5.2 Data Management Methodologies

SMOTE -- Synthetic Minority Over-sampling Technique[9]

SMOTE is a technique for the building of classifiers from datasets that are imbalanced. A dataset is characterized as imbalanced if the classes are not in the same region, i.e. they are not similarly represented in the samples space. Under-sampling of the majority class (i.e. the “normal” category) has been considered in some cases in the literature, as a good way of improving the sensitivity of a classifier compared to the minority class.

In VPH2 we have adopted a mixed method consisting of over-sampling of the minority class (“abnormal” category) on one hand and under-sampling the majority class(“normal” category) on the other, that is able to reach better classifier performance (in ROC space) when compared with the simple, trivial under-sampling of the majority class. This mixed method can also achieve better classifier performance (in ROC space) than varying the loss ratios (in Ripper) or class priors (in Naive Bayes).

The imbalance issue is very important and it hinders knowledge extraction in any data set it appears: Imbalance on the order of 100 to 1 is common in fraud detection and imbalance of up to 100,000 to 1 has been reported in other applications [20]. In another work the SHRINK system was proposed that classifies an overlapping region of minority (positive) and majority (negative) classes as positive (i.e. it adopts the minority class); it searches for the “best positive region" [21].

Other common approaches are:

- “Random re-sampling”, that proposes random re-sampling of the smaller class until it consists of equal number of samples with the majority class.
- “Focused re-sampling”, that proposes re-sampling of the minority examples that occur on the boundary between the two classes.
- “Random under-sampling”, that proposes random under-sampling of the majority class it consists of equal number of samples with the minority class.
- “Focused under-sampling” that proposes under-sampling the majority class samples lying further away from the boundary between the two classes.

One approach that is quite relevant to the one adopted for VPH2 work and thus it is worth referring to it, is the work of Ling and Li [22]. They proposed a combination of over-sampling of the minority class with under-sampling of the majority class and they preferred lift analysis, and not accuracy in order to measure the improvement in a classifier's performance. They further proposed that the test examples can firstly be ranked by confidence and then lift can be used as the evaluation criteria. Solberg and Solberg [23] also considered the problem of imbalanced data sets (in oil slick classification from SAR imagery). They also used over-sampling and under-sampling techniques to improve the classification of oil slicks. To overcome the imbalance problem, they over-sampled (with replacement) 100 samples from the oil slick, and they randomly sampled 100 samples from the non oil slick class to create a new dataset with equal probabilities. Domingos [24] compares the "meta-cost" approach to each of majority under-sampling and minority over-sampling. He finds that meta-cost improves over either, and that under-sampling is preferable to minority over-sampling. Error-based classifiers are made cost-sensitive. A feed-forward neural network trained on an imbalanced dataset may not learn to discriminate enough between classes [25]. The authors proposed that the learning rate of the neural network can be adapted to the statistics of class representation in the data. Lewis and Catlett [26] examined heterogeneous uncertainty sampling for supervised learning. This method is useful for training samples with uncertain classes. The information retrieval (IR) domain [27-30] also faces the problem of class imbalance in the dataset.

The approach adopted in VPH2 proposes an over-sampling of the minority by creating "artificial" instances instead of over-sampling with replacement. This idea is actually inspired by a method that proved very useful in handwritten character recognition [31]. They created additional, artificial training data by modifying, through certain operations, the real data. The operations included rotation and skew that are "natural" ways to change the training data set. In VPH2 artificial instances were generated in a more generic way, by working in "feature space" instead of "data space". The minority class is over-sampled by taking each instance from the minority class and importing artificial instances beside the line segments joining any/all of the k minority class closest neighbors. VPH2 implementation currently uses ten nearest neighbors. For example, if the total over-sampling required is 200%, only two neighbors from the ten nearest neighbors are selected and one sample is generated in the direction of each. Artificial instances are generated as described in the following:

1. Calculate the difference between the feature vector (instance) being considered and its nearest neighbor.
2. Multiply this difference by a randomly chosen number (in the space between 0 and 1), and add it to the feature vector being considered.

This procedure leads to the selection of a random point along the line segment between two particular features and efficiently generalizes the decision region of the minority class.

The artificial instances "force" the classifier to build larger and "vaguer" decision regions, instead of smaller and stricter regions. The minority class instances learn more generalized regions instead of those being learned by the majority class instances in the same region.

The application of SMOTE provides a new aspect of over-sampling. The mixed application of SMOTE and under-sampling seems to be more efficient than plain under-sampling. SMOTE was tested on several datasets and provided improved accuracy compared to other approaches. The mixed application of SMOTE and under-sampling also seems to be more efficient, based on results depicted in ROC space, in comparison with varying loss ratios (in RIPPER) or by varying the class priors (in Naive Bayes); these methods that could straightforwardly handle the skewed class distribution.

Wrapper

Feature selection is the issue of selecting the most pertinent subset of and ignores the rest variables/features that seem to be less important for the classification that must be performed. In order to reach the best possible accuracy/performance with a specific learning algorithm on a specific training data set, the feature selection method must consider the interaction between the algorithm and the training data set.

The adopted in VPH2 project wrapper methodology searches for the best possible feature subset adapted to a certain algorithm and the respective domain. The task of the training algorithm, or the *inducer,* is to induce/ train a *classifier* which will be functional when classifying future cases. What the classifier does is a mapping of features to class values. In the adopted wrapper approach [32], the feature selection method acts as a wrapper around the induction/training algorithm. The feature selection method explores the full data set for a functional subset using the induction/training algorithm as it is and as part of the function that evaluates the possible feature subsets.

The concept behind the wrapper approach is straightforward: the induction/training algorithm is considered to be a black box. It is thus run on the data set, typically split into internal training and holdout sets, with diverse sets of features detached from the data. The feature subset with the maximum evaluation score is selected as the final set and the induction/training algorithm is applied on it. The classifier that is produced is then evaluated based on an independent (holdout) test set that was omitted throughout the training. The purpose of feature subset selection is to find a subset of the original data set, ensuring that whenever an induction/learning algorithm is applied on data including only these selected features builds a classifier with the maximum accuracy. Of course feature selection creates a subset of features choosing from real and existing features, and does not create new features.

The feature selection method looks for a good sub set using the induction/training algorithm itself in the evaluation function. The estimation of the accuracy of the produced classifiers is based on accuracy estimation techniques [33].

The wrapper explores the space of potential parameters. This exploration requires:

- a state space
- an initial state
- a termination condition
- a search engine [34;35]

In the following a short comparison of the two most commonly used search engines is presented: hill-climbing and best-first search. For a total of n features, n bits exist in each state, and every bit indicates if a feature is present (i.e. 1) or absent (i.e. 0). Operators define the connectivity among the states, and operators that append or erase a particular feature from a state are used, equivalent to the search space usually used in stepwise methodologies in Statistics domain.

As an example we assume that we have a state space and operators for a 4 feature problem. The range of the search space for n variables/ features is 0, so it is obviously unreasonable to explore the whole space thoroughly, unless n is small. In the following the different search engines are compared.

The aim of the hill climbing search is to discover the state that is best evaluated, by means of a heuristic function to direct it. Since the accuracy of the induced classifier is still unknown, we can make use of accuracy estimation as both the heuristic and the evaluation functions do. The evaluation function that we employ is 5-fold cross-validation repeated several times. The amount of repetitions is determined per case by looking at the accuracy estimate and the standard deviation that it presents, considering that they are independent. If the standard deviation is higher than 1% and 5 cross-validations still have not been executed, we trigger an additional cross validation run. Despite the fact that this is just a heuristic, it performs well in practice and avoids numerous cross-validation runs in cases of large datasets. This heuristic has the useful characteristic that it executes cross-validation less times on large datasets than on smaller datasets. Since smaller datasets need reduced time to learn, the overall accuracy estimation time, which is the result of the induction/ training algorithm running time and the time needed for cross-validation, grows slower. This way “hardness” is preserved through the use of this heuristic: small data sets are cross-validated several times in order to face the variation that is the result when working with purely populated data sets. For huge datasets, the best approach is to change to a holdout heuristic in order to save more time.

Best-first search [34;35] is more sturdy than hill-climbing. The basic idea is the selection of the most promising node we have constructed to this point and which has not previously been expanded. Best-first search typically terminates when it accomplishes the goal. Since in VPH2 the problem is actually an optimization problem the search can end at any spot and the best solution found hitherto can be returned (supposedly improving over time) at any time making thus the algorithm what is called an “anytime algorithm” [36].

In fact, we should anyway stop the run at some stage, and we employ what is called a ***stale search:*** if an improved node wasn’t found in the previous ***k*** expansions, the search is stopped. An improved node is the node with an accuracy estimation at least E higher than the best one found thus far.

5.3 Data Mining Methodologies

In order to decide which algorithm will be used in the Decision Support System of the VPH^2 platform the following methodologies were applied to Mario Negri data set. Those that proved to be the most useful and efficient in terms of accuracy and transparency were also applied in Niguarda dataset. In any case and for the sake of the deliverable's completeness a short overview of the applied methods is provided in this section while more details are given for the main methods adopted in VPH2 (PART, Decision Trees, Decision Tables, kNN).

Naive Bayes Classifier [37;38]

A Naive Bayes Classifier is a simple probabilistic classifier that estimates the conditional probability of an instance to belong in a specific class using the Bayes theorem. Naive Bayes Classifier assumes that all attributes are conditionally independent given the class. In order for the variable X to be conditionally independent from the variables Y and Z the following condition must be true:

$$P(X \mid Y, Z) = P(X \mid Z)$$

Naïve Bayes Classifier classifies an instance t = {t_1, t_2, …,t_n} to the class the Bayes theorem is applied thus for every value of the class:

$$P(Class = c \mid x_2 = t_2, \ldots, x_n = t_n) = \frac{P(Class = c)\prod_{i=1}^{n} P(x_i = t_i \mid Class = c)}{P(x_1 = t_1, x_2 = t_2, \ldots, x_n = t_n)}$$

The instance is then classified to belong to the class having the probability mentioned above maximum.

Bayesian Network [37;38]

A Bayesian Network is a graphical model that represents the probabilistic relationships between variables. In such a graphical model each vertex is a variable or a group of variables and each edge is the probabilistic relationship between the variables which it connects, for each conditional distribution between variables a direct edge is added.

In Figure 1 a Bayesian Network is depicted and the conditional probabilities *P* (AMI | FollowUp).

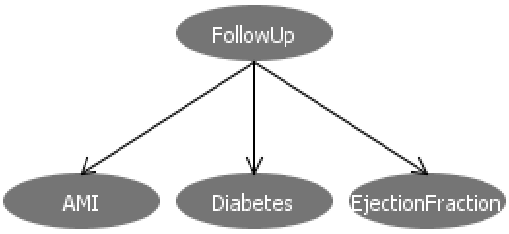


Figure 2: A Bayesian Network and the conditional probabilities *P* (AMI | FollowUp).

The decision making using Bayesian networks is similar to the one using Naive Bayes classifier, considering the "parents" of the of the class vertex.

Multilayer Perceptron [37;38]

Multilayer Perceptron (MLP) is the most successful neural network model in the category of pattern recognition. The Multilayer Perceptron consists of the input layer, the output layer and one or more hidden (intermediary) layers of neurons. The input and output neuron layers generally have linear activation function, contrary to the hidden layers in which neurons have non linear, usually sigmoid functions. In feed – forward multilayer perceptron, each node from each layer is connected with all the nodes from the next. The goal in training a multilayer perceptron is to find the optimal parameters $w_{ij}^{(k)}$, which is the weight of the connection of the neuron j in layer k to neuron i in layer (k+1), and $b_j^{(k)}$, which is the bias of the neuron j in layer k, in order to minimize the total sum of squared errors:

$$E(w) = \frac{1}{2}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$$

Where $\hat{y}_i$ is the output of the multilayer perceptron and y_i the desired output. The algorithm consists of two steps the forward and the backward pass.

- In the forward pass, the outputs corresponding to the inputs are computed.
- In the backward pass the error is propagated backwards through the network and weights are changing using gradient descent.

Radial Basis Function Network [37]

A Radial Basis Function (RBF) network is a neural network that has an input layer, an output layer and in most cases one hidden layer. The activation function of the neurons in a RBF network, is radial basis function the most common function used is a Gaussian transfer function. The concept behind the RBF networks is that instances in a close distance are more possible to have the same predicted value. During RBF network training, one or more neurons are placed in the instances space, in our methodology the position and radius of the neurons (centre and deviation of Gaussian kernel) is decided by applying the algorithm k – nearest neighbors. In training phase the weights of the connection between the neurons are optimized in order to minimize the total sum of squared errors:

$$E(w) = \frac{1}{2}\sum_{i=1}^{N}(y_i - \overset{\wedge}{y_i})^2$$

Where $\overset{\wedge}{y_i}$ is the output of the RBF network and y_i the desired output.

K nearest Neighbours [38;39]

K nearest neighbours classifier is a part of a more general technique called instance based learning. K nearest neighbours does not require building a classification model. In order to classify an instance using K – NN a proximity (distance) measure is required, the distance between the instance to be classified and all the instance of the training set is computed. The k nearest instances are obtained and the class of the instance is decided based on the majority class of its k nearest neighbours.

Voting Feature Intervals [40]

In VFI classification during training the algorithm constructs an interval for each feature, which represents a set of values for the feature, the interval is represented by a vector containing the lower bound, and number of instances from each class that belongs to the specific interval, the upper bound of the interval can be found by checking the lower bound of the next interval. In order to classify a new instance the algorithm checks in which interval each feature of the instance falls, each feature then gives a vote for each class equal to the ratio of the count of the class in the interval to the overall class count. A vector is then constructed for each feature containing the votes for each class. The vectors are then summed up and the predicted class is the one with the highest total vote.

Decision Table [41]

A Decision Table consists of two parts the schema, which is a set of features included in the Decision Table and the body which is a set of labelled instances containing the features described in the body. In order to get the Decision Table rules, given an unlabelled instance the algorithm searches in the data set to find matching instances, the search is done by looking only the features that belong in the schema. The predicted class of the instance is the class of majority of the matched instances, if no instances match the pattern the predicted class is the majority class of the data set. The key for to learning a Decision Table is to select a schema with highly discriminative features.

Decision Table Naive Bayes Combination [42]

The combination of Naive Bayes and Decision Table is a simple Bayesian Network which uses the Decision Table to represent the conditional probabilities. The algorithm for learning the Decision Table - Naive Bayes combination model is similar to the one from learning the Decision Table model described above. At each step in the search of matching instances the algorithm uses an evaluation measure to the best split of the features in two subsets, one for the Decision Table and one for the Bayesian Network.

Repeated Incremental Pruning to Produce Error Reduction (RIPPER)[43]

This algorithm scales almost linearly with the number of training examples and is particularly suited for building models from data sets with imbalanced class distributions. RIPPER also works well with noisy data sets because it uses a validation set to prevent model over fitting. RIPPER chooses the majority class as its default class and learns the rules detecting the minority class. For multi-class problems, the classes are ordered according to the frequencies. Let (y1, y2... yc) be the ordered classes, where y1 is the least frequent class and yc is the most frequent class. During the iteration instances that belong to y1 are labelled as positive examples, while those that belong to other classes are labelled as negative examples. Next, RIPPER extracts rules that distinguish y2 from other remaining classes. This process is repeated until we are left with yc, which is designated as the default class.

Ripper employs the general-to-specific strategy to grow a rule and the FOIL's information gain measure to choose the best conjunct to be added into the rule antecedent. It stops adding conjuncts when the rule starts covering negative examples. The new rule is then pruned based on its performance on the validation set. The following metric is computed to determine whether pruning is needed: (p-n)/(p+n), where p(n) is the number of positive (negative) examples in the validation set covered by the rule. If the metric improves after pruning then the conjunct is removed.

Non Nested Generalised Exemplars (NNGE)[44]

This innovative algorithm generalises exemplars without nesting or overlap. NNGE is the extension of NGE algorithm [45], which generalises by merging exemplars, forming hyperrectangles in feature space that represent conjunctive rules with external disjunction. NNGE forms a generalisation whenever a new instance is imported to the database, by associating it to its closest neighbour of the same class. NGEE does not permit hyperrectangles to overlap or nest.

NNGE algorithm is trained incrementally: it firstly classifies and then generalises each new instance. For that purpose a modified Euclidian distance function is used that handles hyperrectangles, symbolic features, and exemplar and feature weights. Normalisation of numeric feature values is performed by dividing each value by the range of values observed. And the class predicted is the class of the single nearest neighbour. Moreover, NNGE uses dynamic feedback to regulate exemplar and feature weights each time a new instance is classified. During instances' classification, more than one hyperrectangles may be found in which the new instance belongs, but which may be of the mistaken class. NNGE prunes these in order the new instance to no longer be a member.

Once it is classified, the new example is generalized by merging it with the nearest exemplar of the same class, which may be either a single example or a hyperrectangle. In the former case NNGE creates a new hyperrectangle, whereas in the latter it grows the nearest neighbor to encompass the new example. Over-generalization, caused by nesting or overlapping hyperrectangles, is not allowed. Before a new example is generalized, it checks to see if there are any examples in the affected area of feature space that conflict with the proposed new hyperrectangle. If so, the generalization is aborted and the example is stored verbatim.

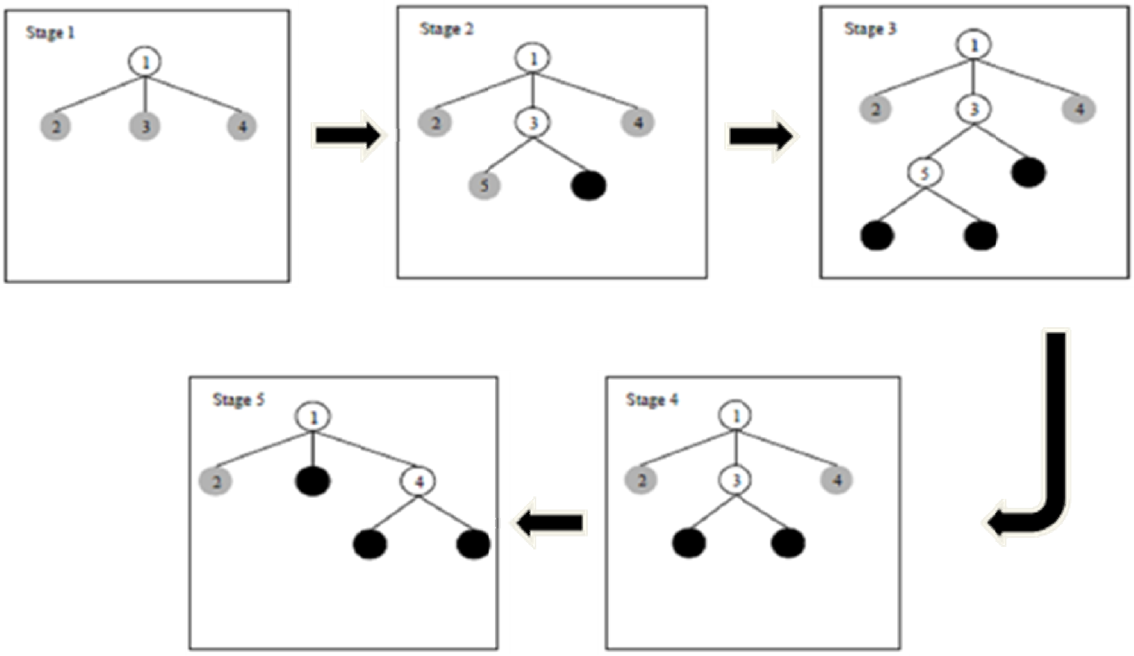
PART[46]

The method is a combination of C4.5 and RIPPER, which are the two most popular schemes for rule learning and which both are adopt a two stages approach. At the first stage they produce a set of rules that they refine at the second. This second stage optimizes the set of rules by either omitting (in C4.5) or by adjusting (in RIPPER) the various rules in order to improve their overall performance and their "cooperation" for producing decisions. PART is a methodology for inferring rules by repetitive generation of partial decision trees. It combines thus the two aforementioned methodologies (C4.5 and RIPPER) by generating rules from decision trees and then by utilizing the "divide and conquer" rule learning method. PART is simple and well-designed. Moreover, tests on standard experimental datasets demonstrate that the resulting rule sets are comparable both in terms of accuracy and in terms of size to those generated when using C 4.5 and are more precise than those generated using RIPPER. PART is a rule induction method that even though it avoids global optimization it generates rule sets that are both accurate and solid. PART has taken its name by partial decision trees in which it is based.
As it is already stated above, PART does not require global optimization to generate its set of rules and this additional straightforwardness is actually its main improvement. By adopting the "divide and conquer" strategy, it first produces a rule, then removes the instances that are covered by this rule and keeps building rules recursively for the residual instances until none is left. Its main difference is the technique based on which a single rule is built:

- a pruned decision tree is built for the current set of instances
- the leaf with the largest coverage is made into a rule
- that tree is discarded

This approach avoids rushed oversimplification by generalizing when and only when the implications are established. The usage of a pruned tree to get a rule, instead of building it incrementally by adding conjunctions sequentially, overcomes the over pruning problem of the fundamental “divide and conquer” rule learner.

The main idea in PART is to construct a partial decision tree and not of a completely explored tree. A partial decision tree is a regular decision tree that has branches to undefined sub-trees. Such a tree is generated by integrating the building and pruning stages with the purpose of finding a stable sub-tree that cannot be further cut down. When this sub-tree has been created, tree building ceases and a single rule is produced. The tree building algorithm is depicted in Figure 1 below.



PART is executed as described in the following: A set of examples is split recursively into a partial tree. A single test is chosen and the examples included are divided into subsets accordingly. The choice is made in exactly the same manner as it is made in C4.5. Then the various subsets are expanded according to their average entropy, starting with the smallest. This procedure continues recursively until a subset is expanded into a leaf and then continues further by backtracking. But as soon as an internal node appears which has all its children expanded into leaves, pruning begins the algorithm checks whether that node is better replaced by a single leaf.

This is just the standard “sub-tree replacement” operation of decision-tree pruning, and the proposed implementation makes the decision in exactly the same way as C4.5. If replacement is performed the algorithm backtracks in the standard way, exploring siblings of the newly-replaced node. However, if during backtracking a node is encountered all of whose children are not leaves and this will happen as soon as a potential sub-tree replacement is not performed then the remaining subsets are left unexplored and the

[47]corresponding sub-trees are left undefined. Due to the recursive structure of the algorithm this event automatically terminates tree generation.

A particular rule is extracted each time a partial tree is built and finalized. Each leaf is potentially a new rule, and the algorithm looks for the "best" leaf of those sub-trees (which are usually the minority) that have been expanded into leaves. PART aims at the most broad rule by selecting the leaf that covers the maximum number of instances.

Missing values in PART are treated in precisely the same manner as in C4.5: if an instance cannot be assigned deterministically to a branch because of a missing attribute value, it is assigned to each of the branches with a weight proportional to the number of training instances going down that branch, normalized by the total number of training instances with known values at the node.

Because a decision tree can be built in time O (an log n) for a dataset with n examples and a attributes, the time taken to generate a rule set of size k is O (kan log n). Assuming (as the analyses of [43;47]] do) that the size of the final theory is constant, the overall time complexity is O (an log n), as compared to O (an log ^{2} n) for RIPPER.

Decision Tree Induction (C 4.5) [48]

Decision tree classifiers are another straightforward and broadly used classification method. Typically each tree has three different types of nodes:

- Root nodes, which do not have any incoming edges any may have 0 or more outgoing edges
- Internal nodes, which have one incoming edge and 2 or more outgoing
- Leaf or terminal nodes, which have one incoming edge and do not have any outgoing

Each leaf of the decision tree is assigned a class label. The root and any other internal nodes contain attribute test conditions to split records that have different features.

The classification of a test record is simple after a decision tree is built. Starting from the root node the test condition is applied to the record and the suitable branch, based on the result of the test, is followed.

Random Forest [49;50]

Random forest is a class of ensemble methods specifically designed for decision tree classifiers. It combines the predictions made by multiple decision trees, where each tree is generated based on the values of an independent set of random vectors. The random vectors are generated from a fixed probability distribution, unlike the adaptive approach used in AdaBoost, where the probability distribution is varied to focus on examples that are hard to classify. Bagging using decision trees is a special case of random forests, where randomness is injected into the model-building process by randomly choosing N samples, with replacement, from the original training set.

5.4 Results

a. Mario Negri Results

In the following chapter the most significant results of the data mining work in Mario Negri data set are presented.

In Table 4 below the results (specificity, sensitivity, accuracy) of the classifiers we have tested with ejection fraction as the targeted outcome are presented. All these classifiers were trained with certain variables (depicted in Table 3 below) indicated by the clinicians involved in this data mining study.

Table 3: Variables from Mario Negri dataset from classification targeting Ejection Fraction

ATTRIBUTES
Gender
Smoke
Nofcigarettes
AMI
Family Diabetes
Family Hypertension
Claudicatio intermittens
B blockers
Ace inhibitor
Calcium channel blockers
Lipid lowering
Diuretics
Wine Intake
PUFA
BMI
Age

Actually the classifiers with Ejection Fraction as outcome were not used in the decision support and it was much more an exercise and example for the collaboration with the clinicians who hadn’t previously been involved in such machine learning studies. Moreover, the accuracy in this study was significantly low and consequently not useful for any decision support.

Table 4: Results classifiers with Ejection Fraction as outcome

Method	Specificity	Sensitivity	Accuracy
Bayes Network Local Search	65.83%	73.18%	69.70%
Bayes Network Global Search	66.41%	73.36%	70.06%
Bayes Network Fixed Search	64.88%	73.88%	69.61%
Bayes Network Ci Search	65.45%	73.53%	69.70%
Naïve Bayes	65.45%	73.53%	69.70%
Naïve Bayes Simple	64.88%	72.32%	68.79%
Naïve Bayes Updateable	64.68%	72.84%	68.97%
Logistic	63.15%	73.01%	68.33%
MLP	54.89%	80.62%	68.43%
RBF Network	62.57%	71.80%	67.42%
Simple Logistic	62.57%	76.30%	69.79%
SMO	62.57%	74.91%	69.06%
Voted Perceptron	10.56%	94.81%	54.87%
IB1	54.89%	64.53%	59.96%
K Nearest Neighbors	61.04%	75.61%	68.70%
k*	57.58%	64.53%	61.24%
LWL	58.35%	70.24%	64.60%
Hyperpipes	99.42%	0.00%	47.13%
Voting Feature Intervals	67.95%	69.90%	68.97%
Conjuctive	50.86%	71.97%	61.97%
Decision Table	60.08%	77.16%	69.06%
DTBN	56.81%	78.89%	68.43%
RIPPER	63.53%	72.15%	68.06%
OneR	64.68%	68.17%	66.52%
PART	61.61%	67.13%	64.51%
Ridor	49.71%	79.58%	65.42%
ADTree	65.07%	72.32%	68.88%
Decision Stump	50.86%	71.97%	61.97%
FT	59.50%	69.72%	64.88%
C 4.5	63.29%	73.53%	68.66%
C 4.5 graft	59.31%	76.64%	68.43%
LAD Tree	62.96%	73.53%	68.52%
LMT	63.92%	74.57%	69.52%
NBTree	65.45%	73.53%	69.70%
Random Forest	62.00%	71.80%	67.15%
Random Tree	60.84%	63.67%	62.33%
RepTree	58.73%	75.61%	67.61%

Next we have started working with late onset heart failure as the targeted outcome. The issue was that the dataset was very unbalanced since there were 101 cases of patients who actually developed late onset heart failure against 1123 who didn’t develop. To overcome this problem the first approach was based in building stratified balanced datasets, i.e. we have created 10 subsets of patients that didn’t develop late onset heart failure each one consisting of 112 patients. Then we have started applying/ testing the various algorithms using as training datasets the resulting 10 balanced training datasets, each one consisting of 213 samples: the 101 patients that developed late onset HF and the 10 different 112 patients' datasets that didn't. The default values of the parameters of the algorithms used are depicted in Table 5.

Table 5: Default Values When Testing Methods.

Method	Parameter	Value
Bayes Network	Conditional probability estimator algorithm	Simple Estimator
	Method for searching network structures	K2
Multilayer Perceptron	Layers	1
	Neurons	(attributes + classes) / 2
	Learning Rate	0.3
	Momentum	0.2
	Number of epochs	500
RBF Network	Minimum cluster's standard	0.1
	K-Means cluster number	2
	Ridge value	10^{-8}
K Nearest Neighbors	K	10
	Nearest neighbour search algorithm	Linear Search
	Distance	Euclidian
Voting Feature Intervals	Bias	0.6
Decision Table	Measure to evaluate the performance attribute subsets	Accuracy
	Search method for attribute subsets	Best First
Decision Table Naive Bayes Combination	Measure to evaluate the performance attribute subsets	Accuracy
	Search method for attribute subsets	Forward selection /backward
RIPPER	Minimum total weight of instances in a rule	2
	Number of optimizations	2
Non Nested Generalised Exemplars	Number of attempts for generalization	5
	Number of folders for mutual information	5
PART	Reduced error prunning	Yes
	Minimum number of instances per rule	2
C 4.5	Reduced error prunning	Yes
	Minimum number of instances per rule	2
Random Forest	Number of randomly chosen attributes.	$\log_2$(number_of_attributes) + 1
	Maximum depth of the tree	Unlimited
	Number of trees	10
Random Tree	Number of randomly chosen attributes.	$\log_2$(number_of_attributes) + 1
	Maximum depth of the tree	Unlimited

The results depicted in Table 7 below are the average of the 10 subsets. Moreover, the variables were split according to the series of patient specific data collection the clinician follows during the assessment of patients' condition on the routine daily practice (Table 6).

Table 6: Variables according to the series the clinician follows for the assessment of patients' condition on the routine daily practice

Demographics	Anthropometrical	Drugs	Physical findings	Biochemical	Echocardiography	Genetics
Gender	BMI	PUFA	SBP	Total Cholesterol	Ejection Fraction	rs4291 allele 1
Smoke		B blockers	DBP	Hdl Cholesterol		rs4291 allele 2
Number of cigarettes		ACE inhibitor	BPM	Triglerides		rs5443 allele 1
Wine intake		Calcium Channel Blockers	Claudicatio intermittens	Fibrinogen		rs5443 allele 2
Age		Lipid lowering	AMI	Aematocrit		rs4646994 allele 1
		Diuretics	Diabetes	White Bloodcell counts		rs4646994 allele 2
			Hypertension	Glycaemia		
				Creatinine		
				UricAcid		
				PCR		
				SGOT		
				SGPT		
				NA		

The five methodologies that are most commonly used in this type of clinical data mining problems were further studied and multiple parameters were tested for each of these methods. In Table 8 the results are depicted.

The second methodology used for balancing the unbalanced dataset was the Synthetic Minority Over – Sampling TEchnique (SMOTE) that is described in chapter 5.2. Using this technique we created balanced dataset, by over – sampling the minority class and creating new instances, for every instance using its ten nearest neighbours. In Table 9 below the specificity, sensitivity and accuracy of the fourteen classifiers that were tested is presented. In order to evaluate the statistical differences of the classifiers with the highest accuracies we have performed the McNemar test. The results are presented in Table 10 below, where NS denotes Non Significant differences, while S denotes significant differences. The McNemar test compares the differences between the classifiers pairwise. In order to compute the statitistical differences between the classifiers the decision differences and agreements between the two classifiers form a contingency table.

	Classifier 2: Positive	Classifier 2: Negative
Classifier 1: Positive	*a*	*b*
Classifier 1: Negative	*c*	*D*

where *a* is the number of instances that both classifier 1 and classifier 2 predict correct, *b* is the number of instances that classifier 1 predicts correct and classifier 2 predict wrong, *c* is the number of instances that classifier 1 predicts wrong and classifier 2 predict correct and *d* is the number of instances that both classifier 1 and classifier 2 predict wrong. In order to compute the McNemar test statistic the following formula is applied:

$$x^2 = \frac{(|b - c| - 0.5)^2}{b + c}$$

Under the null hypothesis, that the two classifiers have no significant statistical differences x^2 has a chi-squared distribution with 1 degree of freedom. If the x^2 result is significant the null hypothesis is rejected, thus the classifiers have significant differences.

In Table 11 the results of the five most commonly used classifiers with various parameter values are shown. As in the methods testing in Table 12 the McNemar tests are depicted in order to notice whether the classifiers have significant differences or not. The classifiers depicted are the ones that had the largest accuracy from each algorithm.

In Table 13 the results of the classifiers are depicted when the features of each dataset are restricted using the Wrapper technique.

Table 7: Results of several methods when the stratified balanced dataset is used

	Demographics Anthropometrical Drugs			Demographics Anthropometrical Drugs Physical Findings			Demographics Anthropometrical Drugs Physical Findings Biochemical			Demographics Anthropometrical Drugs Physical Findings Biochemical Echocardiography			Demographics Anthropometrical Drugs Physical Findings Biochemical Echocardiography Genetics			Genetics		
Method	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
Bayes Network	81.36%	67.13%	74.23%	81.26%	67.13%	74.18%	78.97%	67.23%	73.09%	81.44%	71.39%	76.40%	81.54%	74.75%	78.14%	56.83%	64.16%	60.50%
Naive Bayes	78.67%	70.30%	74.47%	77.77%	69.70%	73.73%	77.68%	62.28%	69.96%	80.35%	63.96%	72.14%	81.25%	64.55%	72.88%	59.90%	55.84%	57.87%
Multilayer Perceptron	68.91%	68.81%	68.87%	68.73%	67.52%	68.13%	72.51%	67.03%	69.76%	76.17%	71.19%	73.68%	78.45%	73.07%	75.76%	66.14%	55.05%	60.59%
RBF Network	75.18%	67.43%	71.30%	73.99%	67.43%	70.70%	76.49%	62.57%	69.52%	77.18%	65.05%	71.10%	77.77%	67.23%	72.49%	58.42%	57.52%	57.97%
K Nearest Neighbours	67.73%	56.93%	62.33%	69.12%	55.84%	62.47%	67.93%	60.00%	63.96%	73.19%	63.37%	68.27%	76.56%	60.69%	68.61%	70.59%	42.77%	56.68%
Voting Feature Intervals	82.25%	65.84%	74.03%	81.85%	65.74%	73.78%	74.89%	67.33%	71.10%	77.86%	71.68%	74.77%	80.94%	70.79%	75.86%	75.45%	43.17%	59.31%
Decision Table	80.49%	66.83%	73.64%	80.49%	66.83%	73.64%	79.79%	65.25%	72.50%	87.01%	69.80%	78.39%	91.46%	71.09%	81.26%	70.10%	48.91%	59.50%
Decision Table Naive Bayes Combination	78.08%	65.45%	71.75%	78.28%	65.84%	72.05%	76.20%	65.64%	70.91%	83.14%	71.88%	77.49%	85.82%	74.65%	80.22%	67.82%	52.57%	60.20%
RIPPER	78.99%	64.46%	71.70%	80.88%	64.16%	72.49%	76.22%	63.37%	69.77%	82.44%	67.43%	74.91%	87.50%	69.01%	78.24%	65.15%	47.13%	56.14%
Non Nested Generalised Exemplars	67.05%	67.33%	67.19%	69.63%	67.92%	68.77%	63.77%	69.80%	66.79%	69.91%	71.58%	70.75%	76.18%	72.67%	74.42%	68.12%	43.47%	55.79%
PART	71.31%	64.75%	68.03%	68.63%	67.62%	68.13%	70.52%	66.63%	68.58%	75.17%	71.78%	73.48%	75.38%	72.97%	74.17%	57.92%	53.66%	55.79%
C 4.5	80.08%	63.66%	71.85%	76.29%	64.55%	70.41%	71.43%	67.03%	69.22%	76.07%	71.98%	74.03%	74.87%	71.39%	73.13%	60.30%	52.38%	56.34%
Random Forest	78.37%	63.07%	70.71%	79.07%	63.56%	71.30%	79.75%	66.04%	72.89%	82.24%	67.23%	74.72%	81.73%	67.82%	74.77%	56.93%	56.73%	56.83%
Random Tree	66.06%	63.56%	64.81%	64.38%	62.97%	63.67%	65.17%	59.90%	62.53%	66.94%	65.84%	66.39%	66.64%	65.45%	66.04%	58.22%	56.24%	57.23%

Table 8: Results of Random forest, c 4.5 part, Multilayer perceptron and Bayes Network using different parameter values and stratified balanced datasets

	Demographics Anthropometrical Drugs			Demographics Anthropometrical Drugs Physical Findings			Demographics Anthropometrical Drugs Physical Findings Biochemical			Demographics Anthropometrical Drugs Physical Findings Biochemical Echocardiography			Demographics Anthropometrical Drugs Physical Findings Biochemical Echocardiography Genetics			Genetics		
Method	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
Random Forest (10 Trees)	78.37%	63.07%	70.71%	79.07%	63.56%	71.30%	79.75%	66.04%	72.89%	82.24%	67.23%	74.72%	81.73%	67.82%	74.77%	56.93%	56.73%	56.83%
Random Forest (20 Trees)	77.48%	63.96%	70.71%	77.58%	64.55%	71.06%	79.17%	66.53%	72.84%	82.53%	69.31%	75.91%	81.73%	69.11%	75.41%	58.12%	57.23%	57.67%
Random Forest (30 Trees)	78.08%	64.36%	71.20%	77.38%	65.54%	71.45%	78.77%	67.33%	73.04%	85.01%	70.40%	77.69%	81.63%	70.79%	76.20%	57.82%	57.82%	57.82%
Random Forest (40 Trees)	77.68%	64.65%	71.16%	78.18%	66.34%	72.24%	78.97%	66.14%	72.54%	84.12%	70.79%	77.44%	82.53%	70.89%	76.70%	57.82%	57.03%	57.43%
Random Forest (50trees)	76.59%	65.64%	71.10%	78.28%	66.04%	72.15%	79.37%	66.93%	73.14%	84.42%	71.09%	77.74%	83.23%	71.09%	77.15%	58.32%	57.62%	57.97%
C 4.5 (Min No Of Instances/Leaf: 2)	82.75%	64.55%	73.64%	81.47%	64.26%	72.84%	80.09%	64.65%	72.35%	86.22%	69.80%	77.99%	86.51%	68.61%	77.55%	60.59%	48.51%	54.55%
C 4.5 (Min No Of Instances/Leaf: 5)	83.24%	64.55%	73.88%	81.66%	63.66%	72.64%	81.07%	62.57%	71.80%	86.81%	67.62%	77.20%	86.52%	68.02%	77.25%	59.11%	45.74%	52.43%
C 4.5 (Min No Of Instances/Leaf: 10)	84.14%	60.40%	72.24%	83.94%	61.19%	72.54%	83.35%	60.40%	71.85%	88.30%	66.04%	77.15%	87.80%	66.14%	76.95%	57.23%	47.23%	52.23%
C 4.5 (Min No Of Instances/Leaf: 20)	77.90%	60.50%	69.17%	78.40%	60.10%	69.22%	80.87%	57.43%	69.12%	85.03%	65.05%	75.01%	83.94%	64.16%	74.02%	59.31%	44.26%	51.78%
Part (Min No Of Instances/Rule: 2)	79.08%	63.17%	71.11%	76.99%	64.85%	70.91%	76.69%	63.47%	70.06%	83.23%	68.22%	75.71%	82.32%	69.80%	76.05%	58.51%	51.49%	55.00%
Part (Min No Of Instances/Rule: 5)	81.15%	64.36%	72.74%	81.16%	63.47%	72.30%	80.46%	62.57%	71.50%	87.10%	65.94%	76.50%	86.19%	68.12%	77.14%	57.72%	48.91%	53.32%
Part (Min No Of Instances/Rule: 10)	80.87%	60.89%	70.86%	81.46%	60.00%	70.71%	81.47%	62.77%	72.09%	88.00%	64.16%	76.05%	86.21%	66.14%	76.15%	57.62%	48.81%	53.22%
Part (Min No Of Instances/Rule: 20)	77.21%	63.37%	70.26%	77.81%	62.48%	70.11%	79.49%	61.68%	70.56%	85.43%	67.13%	76.25%	84.13%	67.62%	75.85%	64.26%	39.01%	51.63%
Decision Table (Search Method: Best First)	80.49%	66.83%	73.64%	80.49%	66.83%	73.64%	79.79%	65.25%	72.50%	87.01%	69.80%	78.39%	91.46%	71.09%	81.26%	70.10%	48.91%	59.50%
Decision Table (Search Method: Greedy Stepwise)	80.78%	66.83%	73.78%	80.78%	66.83%	73.78%	80.09%	65.25%	72.64%	86.92%	69.31%	78.09%	92.06%	70.40%	81.21%	70.69%	50.99%	60.84%
Decision Table (Search Method: Linear Forward Selection)	80.59%	66.44%	73.49%	80.68%	66.14%	73.39%	79.79%	64.75%	72.25%	87.21%	69.41%	78.29%	90.07%	72.18%	81.11%	70.40%	49.31%	59.85%
Decision Table (Search Method: Ranks Search)	80.29%	65.54%	72.89%	80.09%	65.64%	72.84%	79.69%	64.36%	72.00%	84.94%	70.00%	77.45%	86.03%	69.50%	77.74%	75.25%	45.54%	60.40%
Decision Table (Search Method: Scatter Search)	79.79%	67.03%	73.39%	80.19%	66.93%	73.54%	79.10%	66.73%	72.89%	86.72%	69.80%	78.24%	89.48%	70.79%	80.12%	68.81%	52.28%	60.54%

	Demographics Anthropometrical Drugs			Demographics Anthropometrical Drugs Physical Findings			Demographics Anthropometrical Drugs Physical Findings Biochemical			Demographics Anthropometrical Drugs Physical Findings Biochemical Echocardiography			Demographics Anthropometrical Drugs Physical Findings Biochemical Echocardiography Genetics			Genetics		
Method	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
Decision Table (Search Method: Subset Size Forward Selection)	79.89%	66.24%	73.04%	79.20%	66.73%	72.94%	78.90%	65.05%	71.95%	85.04%	69.80%	77.40%	90.47%	70.69%	80.57%	74.06%	47.62%	60.84%
Bayes Network (Method For Searching Network Structures: Ci Search Algorithm)	81.36%	67.13%	74.23%	81.26%	67.13%	74.18%	78.97%	67.23%	73.09%	81.44%	71.39%	76.40%	81.54%	74.75%	78.14%	56.83%	64.16%	60.50%
Bayes Network (Method For Searching Network Structures: Ics Search Algorithm)	80.07%	65.64%	72.84%	79.97%	65.74%	72.84%	78.17%	65.94%	72.05%	81.64%	71.49%	76.55%	81.93%	74.36%	78.13%	66.04%	52.18%	59.11%
Bayes Network (Method For Searching Network Structures: Naive Bayes)	81.36%	67.13%	74.23%	81.26%	67.13%	74.18%	78.97%	67.23%	73.09%	81.44%	71.39%	76.40%	81.54%	74.75%	78.14%	56.83%	64.16%	60.50%
Bayes Network (Method For Searching Network Structures: Global Hill Climber)	81.55%	67.62%	74.58%	81.55%	67.72%	74.62%	79.57%	67.03%	73.29%	82.14%	71.29%	76.70%	82.43%	73.76%	78.09%	57.13%	63.76%	60.45%
Bayes Network (Method For Searching Network Structures: Global K2)	81.36%	67.13%	74.23%	81.26%	67.13%	74.18%	78.97%	67.23%	73.09%	81.44%	71.39%	76.40%	81.54%	74.75%	78.14%	56.83%	64.16%	60.50%
Bayes Network (Method For Searching Network Structures: Global Repeated Hill climber)	81.55%	67.62%	74.58%	81.55%	67.72%	74.62%	79.57%	67.03%	73.29%	82.14%	71.29%	76.70%	82.43%	73.76%	78.09%	57.13%	63.76%	60.45%
Bayes Network (Method For Searching Network Structures: Global Simulated Annealing)	73.11%	64.85%	68.97%	74.90%	67.23%	71.05%	74.70%	66.04%	70.36%	64.40%	54.75%	59.57%	0.00%	0.00%	0.00%	61.68%	61.88%	61.78%
Bayes Network (Method For Searching Network Structures: Global Tabu search)	81.55%	67.62%	74.58%	81.55%	67.72%	74.62%	79.57%	67.03%	73.29%	82.14%	71.29%	76.70%	82.53%	73.86%	78.19%	57.13%	63.76%	60.45%
Bayes Network (Method For Searching Network Structures: Local Hill Climber)	82.94%	64.85%	73.88%	82.84%	64.95%	73.88%	79.67%	64.36%	72.00%	81.75%	72.08%	76.90%	83.72%	73.66%	78.68%	74.26%	38.81%	56.53%
Bayes Network (Method For Searching Network Structures: Local K2)	81.36%	67.13%	74.23%	81.26%	67.13%	74.18%	78.97%	67.23%	73.09%	81.44%	71.39%	76.40%	81.54%	74.75%	78.14%	56.83%	64.16%	60.50%
Bayes Network (Method For Searching Network Structures: Local Lagd Hill Climber)	82.84%	64.85%	73.83%	82.84%	64.95%	73.88%	79.67%	64.36%	72.00%	81.75%	72.08%	76.90%	83.72%	73.66%	78.68%	74.06%	40.10%	57.08%

	Demographics Anthropometrical Drugs			Demographics Anthropometrical Drugs Physical Findings			Demographics Anthropometrical Drugs Physical Findings Biochemical			Demographics Anthropometrical Drugs Physical Findings Biochemical Echocardiography			Demographics Anthropometrical Drugs Physical Findings Biochemical Echocardiography Genetics			Genetics		
Method	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
Bayes Network (Method For Searching Network Structures: Local Repeated Hill climber)	82.94%	64.85%	73.88%	82.84%	64.95%	73.88%	79.67%	64.36%	72.00%	81.75%	72.08%	76.90%	83.72%	73.66%	78.68%	74.26%	38.81%	56.53%
Bayes Network (Method For Searching Network Structures: Local Simulated Annealing)	83.05%	66.63%	74.82%	81.77%	66.44%	74.08%	82.45%	65.25%	73.83%	84.22%	70.79%	77.49%	0.00%	0.00%	0.00%	73.56%	45.54%	59.55%
Bayes Network (Method For Searching Network Structures: Local Tabu search)	81.65%	67.43%	74.53%	81.85%	67.52%	74.67%	79.17%	66.93%	73.04%	82.14%	71.68%	76.90%	82.73%	72.87%	77.79%	78.42%	35.35%	56.88%
Bayes Network (Method For Searching Network Structures: Local Tan)	78.48%	68.61%	73.53%	78.58%	68.51%	73.53%	77.38%	69.21%	73.28%	81.93%	72.28%	77.09%	81.44%	75.15%	78.29%	59.90%	64.16%	62.03%
Multilayer Perceptron (1 Hidden Layer 2 Neurons)	71.91%	65.54%	68.72%	70.73%	67.82%	69.27%	71.02%	66.24%	68.62%	78.64%	69.90%	74.27%	75.86%	73.47%	74.66%	64.65%	52.67%	58.66%
Multilayer Perceptron (1 Hidden Layer Neurons = [No Of Attributes + No Of Classes]/2)	68.91%	68.81%	68.87%	68.73%	67.52%	68.13%	72.51%	67.03%	69.76%	76.17%	71.19%	73.68%	78.45%	73.07%	75.76%	66.14%	55.05%	60.59%
Multilayer Perceptron (1 Hidden Layer Neurons = No Of Attributes)	68.22%	66.63%	67.43%	72.20%	67.72%	69.96%	73.10%	66.73%	69.91%	76.56%	70.50%	73.53%	78.54%	72.57%	75.56%	64.26%	55.05%	59.65%
Multilayer Perceptron (1 Hidden Layer Neurons = No Of Attributes + No Of Classes)	68.83%	66.14%	67.48%	69.92%	68.51%	69.22%	73.50%	67.62%	70.55%	77.36%	70.79%	74.07%	80.03%	72.48%	76.25%	63.37%	55.54%	59.46%

Table 9: Results of several methods using dataset balanced with SMOTE algorithm.

	Demographics Anthropometrical Drugs			Demographics Anthropometrical Drugs Physical Findings			Demographics Anthropometrical Drugs Physical Findings Biochemical			Demographics Anthropometrical Drugs Physical Findings Biochemical Echocardiography			Demographics Anthropometrical Drugs Physical Findings Biochemical Echocardiography Genetics			Genetics		
METHOD	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
Bayes Network	95.19%	87.49%	91.36%	99.02%	90.37%	94.72%	100.00%	90.91%	95.48%	100.00%	90.91%	95.48%	100.00%	90.91%	95.48%	57.17%	89.02%	73.01%
Naive Bayes	84.06%	85.87%	84.96%	83.97%	85.87%	84.92%	74.98%	86.41%	80.66%	78.90%	87.04%	82.95%	82.37%	88.03%	85.18%	62.33%	88.66%	75.43%
Multilayer Perceptron	89.05%	89.02%	89.03%	90.03%	89.38%	89.70%	91.54%	91.18%	91.36%	91.81%	91.63%	91.72%	94.39%	92.17%	93.29%	71.15%	87.58%	79.32%
RBF Network	83.97%	88.39%	86.17%	84.59%	88.12%	86.35%	79.96%	88.03%	83.97%	83.35%	88.48%	85.90%	86.20%	89.56%	87.87%	60.73%	88.75%	74.66%
K Nearest Neighbours	83.97%	89.29%	86.62%	83.97%	90.82%	87.38%	81.75%	91.45%	86.57%	84.59%	91.54%	88.05%	85.49%	92.71%	89.08%	61.89%	88.75%	75.25%
Voting Feature Intervals	84.06%	85.87%	84.96%	83.08%	85.87%	84.47%	84.77%	85.42%	85.09%	87.44%	90.19%	88.81%	88.25%	92.35%	90.29%	50.13%	93.07%	71.49%
Decision Table	93.05%	87.94%	90.51%	97.33%	85.60%	91.50%	99.73%	82.90%	91.36%	99.73%	82.90%	91.36%	99.55%	82.99%	91.32%	71.15%	87.58%	79.32%
Decision Table Naive Bayes Combination	94.57%	89.38%	91.99%	98.40%	90.46%	94.45%	100.00%	90.82%	95.43%	100.00%	90.73%	95.39%	100.00%	90.73%	95.39%	71.15%	87.58%	79.32%
RIPPER	94.84%	81.19%	88.05%	94.48%	83.98%	89.26%	90.20%	84.43%	87.33%	93.14%	86.68%	89.93%	92.52%	87.58%	90.06%	71.15%	87.58%	79.32%
Non Nested Generalised Exemplars	87.62%	87.85%	87.74%	85.66%	88.03%	86.84%	87.36%	77.50%	82.45%	88.42%	77.50%	82.99%	92.52%	79.57%	86.08%	71.15%	60.94%	66.07%
PART	89.31%	89.74%	89.53%	90.12%	89.92%	90.02%	90.74%	88.75%	89.75%	88.25%	89.47%	88.85%	90.92%	91.45%	91.18%	62.24%	88.66%	75.38%
C 4.5	92.43%	88.84%	90.64%	91.45%	89.02%	90.24%	91.45%	88.48%	89.97%	90.83%	88.21%	89.53%	92.25%	91.27%	91.76%	61.71%	88.75%	75.16%
Random Forest	95.37%	89.47%	92.44%	95.81%	89.38%	92.61%	96.79%	90.55%	93.69%	95.90%	91.18%	93.55%	97.95%	91.18%	94.58%	61.89%	88.75%	75.25%
Random Tree	84.95%	87.04%	85.99%	85.22%	88.30%	86.75%	85.40%	86.41%	85.90%	86.46%	88.39%	87.42%	85.93%	88.39%	87.15%	62.15%	88.66%	75.34%

Table 10: Mc Nemar Test of several methods using dataset balanced with SMOTE algorithm

DEMOGRAPHICS ANTHROPOMETRICAL DRUGS							DEMOGRAPHICS ANTHROPOMETRICAL DRUGS PHYSICAL FINDINGS BIOCHEMICAL ECHOCARDIOGRAPHIC						
	Bayes Network	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest		Bayes Network	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest
Bayes Network	NS	S	S	S	S	S	Bayes Network	NS	NS	S	S	S	S
Decision Table	S	NS	S	S	S	S	Decision Table	NS	NS	S	S	S	S
C 4.5	S	S	NS	NS	S	S	C 4.5	S	S	NS	S	S	S
Multilayer Perceptron	S	S	NS	NS	S	S	Multilayer Perceptron	S	S	S	NS	S	S
PART	S	S	S	S	NS	S	PART	S	S	S	S	NS	S
Random Forest	S	S	S	S	S	NS	Random Forest	S	S	S	S	S	NS
DEMOGRAPHICS ANTHROPOMETRICAL DRUGS PHYSICAL FINDINGS							DEMOGRAPHICS ANTHROPOMETRICAL DRUGS PHYSICAL FINDINGS BIOCHEMICAL ECHOCARDIOGRAPHIC GENETICS						
	Bayes Network	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest		Bayes Network	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest
Bayes Network	NS	S	S	S	S	S	Bayes Network	NS	NS	S	S	S	S
Decision Table	S	NS	S	S	S	S	Decision Table	NS	NS	S	S	S	S
C 4.5	S	S	NS	S	S	S	C 4.5	S	S	NS	S	S	S
Multilayer Perceptron	S	S	S	NS	S	S	Multilayer Perceptron	S	S	S	NS	S	S
PART	S	S	S	S	NS	S	PART	S	S	S	S	NS	S
Random Forest	S	S	S	S	S	NS	Random Forest	S	S	S	S	S	NS
DEMOGRAPHICS ANTHROPOMETRICAL DRUGS PHYSICAL FINDINGS BIOCHEMICAL							GENETICS						
	Bayes Network	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest		Bayes Network	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest
Bayes Network	NS	NS	NS	S	S	S	Bayes Network	NS	S	S	S	S	S
Decision Table	NS	NS	NS	S	S	S	Decision Table	S	NS	S	NS	S	S
C 4.5	NS	NS	NS	S	S	S	C 4.5	S	S	NS	S	NS	NS
Multilayer Perceptron	S	S	S	NS	S	S	Multilayer Perceptron	S	NS	S	NS	S	S
PART	S	S	S	S	NS	S	PART	S	S	NS	S	NS	NS
Random Forest	S	S	S	S	S	NS	Random Forest	S	S	NS	S	NS	NS

Table 11 Results of random forest, C 4.5 Part, Multilayer Perceptron and Bayes Network using different parameter values and dataset balanced with SMOTE.

	Demographics Anthropometrical Drugs			Demographics Anthropometrical Drugs Physical Findings			Demographics Anthropometrical Drugs Physical Findings Biochemical			Demographics Anthropometrical Drugs Physical Findings Biochemical Echocardiography			Demographics Anthropometrical Drugs Physical Findings Biochemical Echocardiography Genetics			Genetics -		
METHOD	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
Random Forest (2 Trees))	95.64%	81.91%	88.81%	96.71%	80.92%	88.85%	93.50%	77.50%	85.54%	93.59%	79.21%	86.44%	94.48%	78.40%	86.48%	62.15%	88.66%	75.34%
Random Forest (10 Trees)	95.37%	89.47%	92.44%	95.81%	89.38%	92.61%	96.79%	90.55%	93.69%	95.90%	91.18%	93.55%	97.95%	91.18%	94.58%	61.89%	88.75%	75.25%
Random Forest (20 Trees)	94.92%	89.92%	92.44%	96.44%	90.64%	93.55%	97.68%	91.27%	94.49%	96.97%	92.08%	94.54%	98.22%	91.27%	94.76%	61.98%	88.75%	75.29%
Random Forest (30 Trees)	94.48%	90.28%	92.39%	96.44%	90.64%	93.55%	97.86%	91.45%	94.67%	97.24%	92.26%	94.76%	98.04%	91.36%	94.72%	61.98%	88.75%	75.29%
Random Forest (40 Trees)	95.01%	90.46%	92.75%	96.53%	90.82%	93.69%	97.42%	91.36%	94.40%	97.33%	92.17%	94.76%	98.13%	91.45%	94.81%	61.98%	88.75%	75.29%
Random Forest (50Trees)	95.28%	90.46%	92.88%	96.71%	90.73%	93.73%	97.86%	91.45%	94.67%	97.06%	91.90%	94.49%	98.22%	91.63%	94.94%	61.98%	88.75%	75.29%
C 4.5 (Min Number Of Instances/Leaf: 2)	91.54%	87.94%	89.75%	91.63%	86.68%	89.17%	90.38%	88.21%	89.30%	89.40%	88.75%	89.08%	94.48%	90.19%	92.35%	61.53%	88.75%	75.07%
C 4.5 (Min Number Of Instances/Leaf: 5)	89.85%	87.13%	88.50%	91.27%	86.41%	88.85%	90.29%	88.12%	89.21%	88.42%	88.03%	88.23%	92.97%	90.37%	91.67%	61.35%	88.84%	75.02%
C 4.5 (Min Number Of Instances/Leaf: 10)	89.14%	86.14%	87.65%	90.74%	85.42%	88.09%	89.49%	86.68%	88.09%	88.07%	87.67%	87.87%	90.92%	89.65%	90.29%	61.35%	88.84%	75.02%
C 4.5 (Min Number Of Instances/Leaf: 20)	86.02%	86.86%	86.44%	86.38%	85.24%	85.81%	87.18%	84.97%	86.08%	86.64%	86.59%	86.62%	87.98%	89.38%	88.68%	61.35%	88.84%	75.02%
PART (Min Number Of Instances/Rule: 2)	92.61%	88.21%	90.42%	93.05%	88.30%	90.69%	94.12%	87.76%	90.96%	92.88%	88.03%	90.47%	94.66%	90.82%	92.75%	62.24%	88.66%	75.38%
PART (Min Number Of Instances/Rule: 5)	91.99%	86.50%	89.26%	91.45%	86.41%	88.94%	92.43%	86.68%	89.57%	92.79%	88.84%	90.82%	92.34%	90.73%	91.54%	61.80%	88.75%	75.20%
PART (Min Number Of Instances/Rule: 10)	90.65%	86.41%	88.54%	92.25%	85.06%	88.68%	90.38%	86.68%	88.54%	90.65%	87.76%	89.21%	90.74%	88.66%	89.70%	61.80%	88.75%	75.20%
PART (Min Number Of Instances/Rule: 20)	87.27%	85.78%	86.53%	85.75%	85.51%	85.63%	88.16%	85.06%	86.62%	88.51%	86.41%	87.47%	90.20%	88.39%	89.30%	61.80%	88.75%	75.20%
PART (Min Number Of Instances/Rule: 25)	86.38%	87.94%	87.15%	89.49%	85.60%	87.56%	91.01%	85.42%	88.23%	89.85%	86.59%	88.23%	90.03%	89.02%	89.53%	61.80%	88.75%	75.20%
Decision Table (Search Method: Best first)	93.05%	87.94%	90.51%	97.33%	85.60%	91.50%	99.73%	82.90%	91.36%	99.73%	82.90%	91.36%	99.55%	82.99%	91.32%	71.15%	87.58%	79.32%

	Demographics Anthropometrical Drugs			Demographics Anthropometrical Drugs Physical Findings			Demographics Anthropometrical Drugs Physical Findings Biochemical			Demographics Anthropometrical Drugs Physical Findings Biochemical Echocardiography			Demographics Anthropometrical Drugs Physical Findings Biochemical Echocardiography Genetics			Genetics -		
METHOD	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
Decision Table (Search Method: Greedy stepwise)	93.50%	87.94%	90.73%	96.97%	85.51%	91.27%	99.73%	82.90%	91.36%	99.73%	82.90%	91.36%	99.73%	82.90%	91.36%	68.21%	87.67%	77.89%
Decision Table (Search Method: Linear forward selection)	93.05%	87.94%	90.51%	97.33%	85.60%	91.50%	99.64%	82.54%	91.14%	99.64%	82.54%	91.14%	99.47%	82.63%	91.09%	71.15%	87.58%	79.32%
Decision Table (Search Method: Rank search)	93.77%	87.67%	90.73%	97.68%	82.45%	90.11%	95.28%	86.14%	90.73%	95.46%	85.96%	90.73%	95.46%	85.87%	90.69%	71.15%	87.58%	79.32%
Decision Table (Search Method: Scattersearchv1)	92.97%	87.85%	90.42%	95.99%	84.34%	90.20%	99.20%	82.72%	91.00%	99.55%	83.26%	91.45%	98.93%	83.17%	91.09%	62.51%	87.94%	75.16%
Decision Table (Search Method: Subsetsize forward selection)	93.14%	87.85%	90.51%	96.97%	85.15%	91.09%	99.55%	82.09%	90.87%	99.55%	82.09%	90.87%	99.73%	82.09%	90.96%	67.14%	87.85%	77.44%
Bayes Network (Method For Searching Network Structures: Ci search algorithm)	95.19%	87.49%	91.36%	99.02%	90.37%	94.72%	100.00%	90.91%	95.48%	100.00%	90.91%	95.48%	100.00%	90.91%	95.48%	57.17%	89.02%	73.01%
Bayes Network (Method For Searching Network Structures: Ics search algorithm)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	57.17%	88.93%	72.96%
Bayes Network (Method For Searching Network Structures: Naive Bayes)	95.19%	87.49%	91.36%	99.02%	90.37%	94.72%	100.00%	90.91%	95.48%	100.00%	90.91%	95.48%	100.00%	90.91%	95.48%	57.17%	89.02%	73.01%
Bayes Network (Method For Searching Network Structures: Global hill climber)	95.01%	87.67%	91.36%	98.84%	90.28%	94.58%	100.00%	90.91%	95.48%	99.91%	90.91%	95.43%	99.91%	90.91%	95.43%	57.17%	89.02%	73.01%
Bayes Network (Method For Searching Network Structures: Global k2)	95.19%	87.49%	91.36%	99.02%	90.37%	94.72%	100.00%	90.91%	95.48%	100.00%	90.91%	95.48%	100.00%	90.91%	95.48%	57.17%	89.02%	73.01%
Bayes Network (Method For Searching Network Structures: Global repeated hill climber)	95.01%	87.67%	91.36%	98.84%	90.28%	94.58%	100.00%	90.91%	95.48%	99.91%	90.91%	95.43%	99.91%	90.91%	95.43%	57.17%	89.02%	73.01%
Bayes Network (Method For Searching Network Structures: Global simulated annealing)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	61.44%	88.75%	75.02%
Bayes Network (Method For Searching Network Structures: Global Tabu search)	95.01%	87.67%	91.36%	98.84%	90.28%	94.58%	100.00%	90.91%	95.48%	99.91%	90.91%	95.43%	99.91%	90.91%	95.43%	57.17%	89.02%	73.01%
Bayes Network (Method For Searching Network Structures:	95.19%	87.49%	91.36%	99.02%	90.37%	94.72%	100.00%	90.91%	95.48%	100.00%	90.91%	95.48%	100.00%	90.91%	95.48%	57.17%	89.02%	73.01%

	Demographics Anthropometrical Drugs			Demographics Anthropometrical Drugs Physical Findings			Demographics Anthropometrical Drugs Physical Findings Biochemical			Demographics Anthropometrical Drugs Physical Findings Biochemical Echocardiography			Demographics Anthropometrical Drugs Physical Findings Biochemical Echocardiography Genetics			Genetics -		
METHOD	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
Local hill climber)																		
Bayes Network (Method For Searching Network Structures: Lk2)	95.19%	87.49%	91.36%	99.02%	90.37%	94.72%	100.00%	90.91%	95.48%	100.00%	90.91%	95.48%	100.00%	90.91%	95.48%	57.17%	89.02%	73.01%
Bayes Network (Method For Searching Network Structures: Local lagd hill climber)	95.46%	88.93%	92.21%	99.20%	90.37%	94.81%	100.00%	90.91%	95.48%	99.91%	90.91%	95.43%	100.00%	90.91%	95.48%	56.63%	89.11%	72.78%
Bayes Network (Method For Searching Network Structures: Local repeated hill climber)	95.19%	87.49%	91.36%	99.02%	90.37%	94.72%	100.00%	90.91%	95.48%	100.00%	90.91%	95.48%	100.00%	90.91%	95.48%	57.17%	89.02%	73.01%
Bayes Network (Method For Searching Network Structures: Local simulated annealing)	0.00%	0.00%	0.00%	99.38%	90.82%	95.12%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	60.91%	88.84%	74.80%
Bayes Network (Method For Searching Network Structures: Local tabu search)	95.10%	87.49%	91.32%	98.84%	90.37%	94.63%	100.00%	90.91%	95.48%	100.00%	90.91%	95.48%	100.00%	90.91%	95.48%	56.72%	89.11%	72.83%
Bayes Network (Method For Searching Network Structures: Local Tan)	96.88%	88.21%	92.57%	99.29%	90.64%	94.99%	100.00%	90.91%	95.48%	100.00%	90.91%	95.48%	100.00%	90.91%	95.48%	61.44%	88.75%	75.02%
Multilayer Perceptron (1 Hidden Layer Neurons = [Number Of Attributes + Number Of Classes]/2)	89.05%	89.02%	89.03%	90.03%	89.38%	89.70%	91.54%	91.18%	91.36%	91.81%	91.63%	91.72%	94.39%	92.17%	93.29%	71.15%	87.58%	79.32%
Multilayer Perceptron (1 Hidden Layer Neurons = Number Of Attributes)	90.03%	87.94%	88.99%	89.31%	90.01%	89.66%	91.81%	91.81%	91.81%	93.05%	92.44%	92.75%	95.19%	92.62%	93.91%	71.15%	87.58%	79.32%
Multilayer Perceptron (1 Hidden Layer Neurons = Number Of Attributes + Number Of Classes)	89.76%	88.30%	89.03%	89.49%	89.20%	89.35%	92.52%	91.36%	91.94%	92.61%	92.17%	92.39%	94.57%	91.90%	93.24%	71.15%	87.58%	79.32%
Multilayer Perceptron (1 Hidden Layer 2 Neurons)	87.36%	88.12%	87.74%	86.20%	89.38%	87.78%	89.67%	89.38%	89.53%	90.56%	90.64%	90.60%	94.03%	91.18%	92.61%	70.35%	88.21%	79.23%

Table 12: Mc Nemar Test of random forest, c 4.5 part, multilayer perceptron and bayes network with using different parameter values and dataset balanced with SMOTE.

DEMOGRAPHICS ANTHROPOMETRICAL DRUGS

	Bayes Network	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest
Bayes Network	NS	NS	NS	S	NS	S
Decision Table	NS	NS	NS	S	S	S
C 4.5	NS	NS	NS	S	S	S
Multilayer Perceptron	S	S	S	NS	S	S
PART	NS	S	S	S	NS	S
Random Forest	S	S	S	S	S	NS

DEMOGRAPHICS ANTHROPOMETRICAL DRUGS PHYSICAL FINDINGS BIOCHEMICAL ECHOCARDIOGRAPHIC

	Bayes Network	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest
Bayes Network	NS	S	S	S	S	S
Decision Table	S	NS	NS	S	NS	S
C 4.5	S	NS	NS	S	NS	S
Multilayer Perceptron	S	S	S	NS	S	S
PART	S	NS	NS	S	NS	S
Random Forest	S	S	S	S	S	NS

DEMOGRAPHICS ANTHROPOMETRICAL DRUGS PHYSICAL FINDINGS

	Bayes Network	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest
Bayes Network	NS	S	S	S	S	S
Decision Table	S	NS	S	S	NS	S
C 4.5	S	S	NS	S	S	S
Multilayer Perceptron	S	S	S	NS	S	S
PART	S	NS	S	S	NS	S
Random Forest	S	S	S	S	S	NS

DEMOGRAPHICS ANTHROPOMETRICAL DRUGS PHYSICAL FINDINGS BIOCHEMICAL ECHOCARDIOGRAPHIC GENETICS

	Bayes Network	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest
Bayes Network	NS	NS	S	S	S	S
Decision Table	NS	NS	S	S	S	S
C 4.5	S	S	NS	S	NS	S
Multilayer Perceptron	S	S	S	NS	S	S
PART	S	S	NS	S	NS	S
Random Forest	S	S	S	S	S	NS

DEMOGRAPHICS ANTHROPOMETRICAL DRUGS PHYSICAL FINDINGS BIOCHEMICAL

	Bayes Network	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest
Bayes Network	NS	NS	S	S	NS	S
Decision Table	NS	NS	S	S	NS	S
C 4.5	S	S	NS	S	S	S
Multilayer Perceptron	S	S	S	NS	S	S
PART	NS	NS	S	S	NS	S
Random Forest	S	S	S	S	S	NS

GENETICS

	Bayes Network	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest
Bayes Network	NS	S	S	S	S	S
Decision Table	S	NS	S	NS	S	S
C 4.5	S	S	NS	S	NS	NS
Multilayer Perceptron	S	NS	S	NS	S	S
PART	S	S	NS	S	NS	NS
Random Forest	S	S	NS	S	NS	NS

Table 13: Results of random forest, c 4.5, Part and bayes network with using different parameter values and dataset balanced with SMOTE using wrapper

	DEMOGRAPHICS ANTHROPOMETRICAL DRUGS			DEMOGRAPHICS ANTHROPOMETRICAL DRUGS PHYSICAL FINDINGS			DEMOGRAPHICS ANTHROPOMETRICAL DRUGS PHYSICAL FINDINGS BIOCHEMICAL			DEMOGRAPHICS ANTHROPOMETRICAL DRUGS PHYSICAL FINDINGS BIOCHEMICAL ECHOCARDIOGRAPHIC			DEMOGRAPHICS ANTHROPOMETRICAL DRUGS PHYSICAL FINDINGS BIOCHEMICAL ECHOCARDIOGRAPHIC GENETICS			GENETICS		
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy
Wrapper C 4.5	83.45%	61.19%	72.30%	82.86%	61.49%	72.15%	82.66%	61.58%	72.10%	86.72%	67.62%	77.15%	87.90%	68.71%	78.29%	61.78%	44.95%	53.37%
Wrapper Decision Table	82.37%	63.17%	72.74%	82.86%	63.47%	73.14%	82.57%	63.17%	72.84%	86.43%	69.70%	78.04%	91.96%	70.00%	80.96%	72.38%	48.81%	60.59%
Wrapper Part	83.95%	59.90%	71.90%	83.06%	60.30%	71.65%	83.35%	60.99%	72.15%	87.81%	67.03%	77.40%	88.70%	67.92%	78.29%	60.69%	46.44%	53.56%
Wrapper Bayes	81.67%	64.95%	73.29%	81.57%	64.95%	73.24%	80.68%	65.54%	73.09%	84.34%	72.08%	78.19%	86.22%	73.66%	79.92%	62.28%	57.62%	59.95%
Wrapper RF	81.96%	62.28%	72.10%	80.46%	61.29%	70.86%	77.09%	64.06%	70.56%	82.73%	68.12%	75.41%	81.94%	69.31%	75.61%	61.29%	55.45%	58.37%

The advantages and disadvantages of each classifier were explained to the clinicians. After the results and the classifiers' outputs were presented to the users it was requested that the decision made by the classifier must be presented clearly (transparently), thus a rule based classifier must be adopted. The clinicians reviewed the rules produced by the rule based classifiers (PART, RIPPER, Decision Table, C 4.5, Random Tree and Random Forest). One more remark was that not all rules were reasonable from a medical point of view and it was decided to use a classifier that could be edited (permit the user to delete rules that were not correct and to add new rules that were common knowledge). Since trees cannot be edited (deleting a leaf will lead to unclassified instances and addition of a new leaf can result a non tree classifier) the choice was taken between PART and RIPPER.

The results of the unbalanced dataset were poor in both accuracy and sensitivity; the classifiers only predicted the patients that didn't develop late onset HF. The stratified balanced datasets yield better results in both accuracy and specificity, but rules extracted from those dataset didn't agree with common knowledge, mainly because the datasets were not large enough.

The results of the classifiers that were built using SMOTE datasets were both more accurate and predicted patients with late onset heart failure; still, the disadvantage was that the rules were in conflict with common knowledge, because there were a lot of features/ variables in each dataset.

The results that were produced with the dataset that the Wrapper technique was applied were also poor in sensitivity.

Our next step, in order to overcome the issues mentioned above and improve the accuracy of the algorithms, was to restrict the dataset to fewer features; those restricted datasets were provided by the doctors. Diabetes, ejection fraction and AMI site that are proven according to literature to be good predictors for late onset heart failure were indicated as the first dataset. Doctors also needed to know how biochemical data, genetics data and PUFA treatment could improve the accuracy in prediction when used in addition with Diabetes, ejection fraction and AMI site, so three more restricted datasets were constructed referred as "Diabetes Ejection Fraction AMI Biochemical", "Diabetes Ejection Fraction AMI Genetics", "Diabetes Ejection Fraction AMI PUFA" in the tables below.

Clinicians also proposed that it would be interesting to see the accuracy of predicting late onset heart failure on patients that are difficult to prognose such as young patients, non diabetic and having ejection fraction larger than 40%, using the genetics features. Accordingly, three more datasets were constructed:

- Genetics for patients with Ejection Fraction over 40%
- Genetics for non diabetic patients
- Genetics for female patients younger than 60 years old or male patients younger than 55 years old.

In Table 14 and Table 15 the results of the datasets restricted by the clinicians and the datasets that include the genetics respectively, are depicted and several algorithms are presented; the results are the average of the corresponding values of the stratified balanced datasets.

In Table 16 and Table 17 the results of algorithms most commonly used in such datasets are shown; the results are the average of the corresponding values of the stratified balanced datasets.

Table 18 and Table 20 show the results of several methodologies when the datasets are balanced using SMOTE. Table 18 contains the datasets with Diabetes, ejection fraction AMI and the added features in order to test the improvement in accuracy and Table 20 contains the results for the datasets including genetics. Tables 19 and 21 contain the McNemar tests for the abovementioned datasets using the best classifiers.

In Tables 22 and 24 the results of the five most commonly used classifiers are depicted. Table 22 contains the results of “Diabetes Ejection Fraction AMI Biochemical”, “Diabetes Ejection Fraction AMI Genetics”, and “Diabetes Ejection Fraction AMI PUFA” datasets. Table 24 shows the results of “Genetics when Ejection Fraction > 40”, “Genetics when Non Diabetic», «Genetics when gender is female and age < 60 or gender is male and gender < 55” datasets. In Table 23 and Table 25 the McNemar tests used to compute whether the classifiers have significant or not significant differences for the abovementioned datasets are depicted.

Table 14: Results of several methods from datasets restricted by clinicians' feedback using stratified balanced datasets

	Diabetes Ejection Fraction AMI			Diabetes Ejection Fraction AMI Biochemical				Diabetes Ejection Fraction AMI Genetics				Diabetes Ejection Fraction AMI PUFA			
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	Improvement of accuracy	specificity	sensitivity	accuracy	Improvement of accuracy	specificity	sensitivity	accuracy	Improvement of accuracy
Bayes Network	59.22%	67.42%	63.29%	100.00%	90.82%	95.43%	32.14%	72.84%	73.81%	73.32%	10.03%	57.97%	68.95%	63.43%	0.13%
Naive Bayes	59.13%	68.95%	64.01%	69.72%	88.93%	79.27%	15.26%	69.81%	73.72%	71.75%	7.74%	57.44%	71.11%	64.23%	0.22%
Multilayer Perceptron	58.42%	72.46%	65.40%	81.75%	89.38%	85.54%	20.14%	76.05%	90.19%	83.08%	17.68%	51.56%	86.59%	68.98%	3.58%
RBF Network	60.28%	69.13%	64.68%	75.16%	86.68%	80.89%	16.20%	69.90%	70.66%	70.28%	5.60%	56.90%	70.66%	63.74%	-0.94%
K Nearest Neighbors	59.84%	69.94%	64.86%	75.33%	89.56%	82.41%	17.55%	67.14%	91.81%	79.41%	14.55%	52.72%	85.60%	69.07%	4.21%
Voting Feature Intervals	58.06%	69.49%	63.74%	79.79%	93.34%	86.53%	22.78%	54.59%	89.02%	71.71%	7.97%	56.72%	72.91%	64.77%	1.03%
Decision Table	61.71%	67.87%	64.77%	98.13%	84.61%	91.41%	26.63%	77.29%	88.21%	82.72%	17.95%	54.23%	85.42%	69.74%	4.97%
Decision Table Naive Bayes Combination	59.31%	70.39%	64.82%	99.91%	90.01%	94.99%	30.17%	77.38%	88.03%	82.68%	17.86%	54.23%	85.42%	69.74%	4.92%
RIPPER	63.22%	68.86%	66.03%	83.08%	89.74%	86.39%	20.37%	75.96%	85.33%	80.62%	14.59%	56.37%	80.47%	68.35%	2.33%
Non Nested Generalised Exemplars	63.31%	56.44%	59.89%	85.31%	80.56%	82.95%	23.05%	73.73%	80.02%	76.86%	16.97%	61.80%	56.98%	59.40%	-0.49%
PART	60.37%	69.67%	65.00%	82.01%	90.82%	86.39%	21.40%	73.82%	87.13%	80.44%	15.44%	52.27%	86.68%	69.38%	4.39%
C 4.5	60.28%	69.58%	64.91%	84.77%	88.48%	86.62%	21.71%	73.02%	87.94%	80.44%	15.53%	52.54%	86.68%	69.52%	4.61%
Random Forest	61.09%	69.22%	65.13%	87.27%	91.45%	89.35%	24.22%	74.44%	87.49%	80.93%	15.80%	53.25%	85.60%	69.34%	4.21%
Random Tree	60.64%	69.49%	65.04%	83.53%	83.17%	83.35%	18.31%	74.71%	87.76%	81.20%	16.16%	53.34%	85.60%	69.38%	4.34%

Table 15: Results of several methods from datasets restricted by clinicians' feedback feedback including genetics using stratified balanced datasets

	Genetics when Ejection Fraction > 40			Genetics when Non Diabetic			Genetics when gender is female and age < 60 or gender is male and gender < 55		
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy
Bayes Network	71.62%	62.50%	66.89%	64.81%	75.60%	70.42%	86.21%	95.72%	90.85%
Naive Bayes	58.29%	64.72%	61.63%	56.14%	78.35%	67.68%	88.40%	92.43%	90.37%
Multilayer Perceptron	56.07%	80.54%	68.78%	65.59%	81.82%	74.02%	95.30%	95.39%	95.35%
RBF Network	65.81%	67.88%	66.89%	62.74%	74.28%	68.74%	88.40%	93.42%	90.85%
K Nearest Neighbors	38.29%	85.92%	63.02%	46.83%	93.42%	71.04%	90.91%	93.09%	91.97%
Voting Feature Intervals	37.78%	86.23%	62.94%	38.68%	94.50%	67.68%	86.52%	95.72%	91.01%
Decision Table	50.09%	84.02%	67.71%	57.96%	89.95%	74.58%	94.36%	95.39%	94.86%
Decision Table Naive Bayes Combination	54.36%	82.59%	69.02%	58.21%	89.95%	74.70%	94.36%	95.39%	94.86%
RIPPER	45.81%	67.56%	57.11%	45.41%	92.94%	70.11%	94.67%	94.41%	94.54%
Non Nested Generalised Exemplars	61.37%	56.96%	59.08%	60.41%	68.54%	64.64%	95.92%	93.42%	94.70%
PART	53.33%	73.89%	64.01%	51.36%	91.63%	72.28%	91.22%	95.07%	93.10%
C 4.5	51.11%	66.14%	58.92%	52.91%	91.27%	72.84%	90.28%	96.05%	93.10%
Random Forest	53.16%	70.73%	62.28%	49.16%	92.58%	71.72%	90.60%	88.82%	89.73%
Random Tree	54.87%	68.99%	62.20%	50.19%	92.22%	72.03%	89.34%	89.14%	89.25%

Table 16: Results of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from datasets restricted by clinicians' feedback using stratified balanced datasets

	Diabetes Ejection Fraction AMI			Diabetes Ejection Fraction AMI Biochemical				Diabetes Ejection Fraction AMI Genetics				Diabetes Ejection Fraction AMI PUFA			
METHOD	specificity	sensitivity	accurac y	specificity	sensitivity	accuracy	Improvement of accuracy	specificity	sensitivity	accuracy	Improvement of accuracy	specificity	sensitivity	accuracy	Improvement of accuracy
Random Forest (10 Trees)	61.09%	69.22%	65.13%	87.27%	91.45%	89.35%	24.22%	74.44%	87.49%	80.93%	15.80%	53.25%	85.60%	69.34%	4.21%
Random Forest (20 Trees)	59.39%	69.94%	64.64%	86.55%	92.26%	89.39%	24.75%	74.35%	87.40%	80.84%	16.20%	53.25%	85.60%	69.34%	4.70%
Random Forest (30 Trees)	59.39%	69.94%	64.64%	86.55%	92.62%	89.57%	24.93%	74.71%	87.13%	80.89%	16.25%	53.25%	85.51%	69.29%	4.66%
Random Forest (40 Trees)	59.39%	69.94%	64.64%	86.73%	92.62%	89.66%	25.02%	74.62%	87.40%	80.98%	16.34%	53.25%	85.51%	69.29%	4.66%
Random Forest (50Trees)	59.39%	69.94%	64.64%	86.38%	92.89%	89.62%	24.98%	74.80%	86.86%	80.80%	16.16%	53.25%	85.51%	69.29%	4.66%
C 4.5 (min number of instances/leaf: 2)	62.60%	67.69%	65.13%	82.28%	91.36%	86.80%	21.67%	72.40%	88.66%	80.48%	15.35%	53.43%	84.16%	68.71%	3.58%
C 4.5 (min number of instances/leaf: 5)	62.60%	67.69%	65.13%	81.75%	91.36%	86.53%	21.40%	70.26%	88.66%	79.41%	14.28%	53.43%	84.16%	68.71%	3.58%
C 4.5 (min number of instances/leaf: 10)	62.60%	67.69%	65.13%	81.48%	91.45%	86.44%	21.31%	67.94%	86.14%	76.99%	11.86%	52.81%	84.16%	68.40%	3.27%
C 4.5 (min number of instances/leaf: 15)	62.60%	67.69%	65.13%	80.85%	90.01%	85.41%	20.28%	68.30%	83.26%	75.74%	10.61%	53.43%	83.53%	68.40%	3.27%
C 4.5 (min number of instances/leaf: 20)	62.60%	67.69%	65.13%	79.25%	91.63%	85.41%	20.28%	69.46%	79.93%	74.66%	9.53%	52.89%	83.53%	68.13%	3.00%
PART (min number of instances/rule: 2)	59.13%	69.94%	64.50%	81.21%	90.46%	85.81%	21.31%	72.57%	87.13%	79.81%	15.31%	52.36%	86.50%	69.34%	4.83%
PART (min number of instances/rule: 5)	59.13%	69.94%	64.50%	82.64%	89.56%	86.08%	21.58%	71.77%	85.87%	78.78%	14.28%	52.36%	86.59%	69.38%	4.88%
PART (min number of instances/rule: 10)	59.13%	69.94%	64.50%	81.39%	91.99%	86.66%	22.16%	70.17%	83.71%	76.90%	12.40%	53.52%	84.52%	68.93%	4.43%
PART (min number of instances/rule:15)	59.13%	69.94%	64.50%	80.23%	91.18%	85.68%	21.17%	71.42%	81.55%	76.45%	11.95%	56.81%	79.57%	68.13%	3.63%
PART (min number of instances/rule: 20)	59.13%	69.94%	64.50%	80.68%	92.17%	86.39%	21.89%	67.59%	82.36%	74.93%	10.43%	56.37%	78.94%	67.59%	3.09%
Decision Table (search method: Best First)	61.71%	67.87%	64.77%	98.13%	84.61%	91.41%	26.63%	77.29%	88.21%	82.72%	17.95%	54.23%	85.42%	69.74%	4.97%
Decision Table (search method: Greedy Stepwise)	61.71%	67.87%	64.77%	98.58%	83.80%	91.23%	26.45%	77.29%	88.21%	82.72%	17.95%	55.74%	82.72%	69.16%	4.39%

	Diabetes Ejection Fraction AMI			Diabetes Ejection Fraction AMI Biochemical				Diabetes Ejection Fraction AMI Genetics				Diabetes Ejection Fraction AMI PUFA			
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	Improvement of accuracy	specificity	sensitivity	accuracy	Improvement of accuracy	specificity	sensitivity	accuracy	Improvement of accuracy
Decision Table (search method: Linear Forward Selection)	61.71%	67.87%	64.77%	98.13%	84.61%	91.41%	26.63%	77.29%	88.21%	82.72%	17.95%	54.23%	85.42%	69.74%	4.97%
Decision Table (search method: Rank Search)	60.64%	68.86%	64.73%	93.59%	83.98%	88.81%	24.08%	77.29%	88.21%	82.72%	17.99%	54.23%	85.42%	69.74%	5.01%
Decision Table (search method: ScatterSearchV1)	60.64%	68.86%	64.73%	98.93%	83.17%	91.09%	26.37%	77.29%	88.21%	82.72%	17.99%	54.23%	85.42%	69.74%	5.01%
Decision Table (search method: Subset Size Forward Selection)	61.89%	69.31%	65.58%	99.91%	82.90%	91.45%	25.87%	77.29%	88.21%	82.72%	17.14%	54.23%	85.42%	69.74%	4.16%
Bayes Network (method for searching network structures: ICS Search Algorithm)	61.89%	68.68%	65.26%	0.00%	0.00%	0.00%	-65.26%	71.33%	82.00%	76.63%	11.37%	57.08%	76.87%	66.92%	1.66%
Bayes Network (method for searching network structures: Naive Bayes)	59.22%	67.42%	63.29%	100.00%	90.82%	95.43%	32.14%	72.84%	73.81%	73.32%	10.03%	57.97%	68.95%	63.43%	0.13%
Bayes Network (method for searching network structures: Global Hill Climber)	59.22%	67.42%	63.29%	100.00%	90.82%	95.43%	32.14%	72.84%	73.81%	73.32%	10.03%	57.97%	68.95%	63.43%	0.13%
Bayes Network (method for searching network structures: gK2)	59.22%	67.42%	63.29%	100.00%	90.82%	95.43%	32.14%	72.84%	73.81%	73.32%	10.03%	57.97%	68.95%	63.43%	0.13%
Bayes Network (method for searching network structures: Global Repeated Hill Climber)	59.22%	67.42%	63.29%	100.00%	90.82%	95.43%	32.14%	72.84%	73.81%	73.32%	10.03%	57.97%	68.95%	63.43%	0.13%
Bayes Network (method for searching network structures: Global Tabu Search)	59.22%	67.42%	63.29%	100.00%	90.82%	95.43%	32.14%	72.84%	73.81%	73.32%	10.03%	57.97%	68.95%	63.43%	0.13%
Bayes Network (method for searching network structures: Local Hill Climber)	59.22%	67.42%	63.29%	100.00%	90.82%	95.43%	32.14%	75.42%	67.60%	71.53%	8.24%	59.22%	67.42%	63.29%	0.00%
Bayes Network (method for searching network structures: lK2)	59.22%	67.42%	63.29%	100.00%	90.82%	95.43%	32.14%	72.84%	73.81%	73.32%	10.03%	57.97%	68.95%	63.43%	0.13%
Bayes Network (method for searching network structures: Local LAGD Hill Climber)	45.59%	79.21%	62.31%	100.00%	90.82%	95.43%	33.12%	85.04%	50.95%	68.08%	5.77%	45.59%	79.21%	62.31%	0.00%
Bayes Network (method for searching network structures: Local Repeated Hill Climber)	59.22%	67.42%	63.29%	100.00%	90.82%	95.43%	32.14%	75.42%	67.60%	71.53%	8.24%	59.22%	67.42%	63.29%	0.00%
Bayes Network (method for searching network structures: Local Tabu Search)	45.59%	79.21%	62.31%	100.00%	90.82%	95.43%	33.12%	75.42%	66.79%	71.13%	8.82%	52.63%	71.74%	62.13%	-0.18%
Bayes Network (method for searching network structures: Local TAN)	65.36%	60.67%	63.03%	100.00%	90.82%	95.43%	32.41%	72.22%	75.79%	73.99%	10.97%	58.95%	69.58%	64.23%	1.21%

Table 17: Results of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from datasets restricted by clinicians' feedback including genetics using stratified balanced datasets

	Genetics when Ejection Fraction > 40			Genetics when Non Diabetic			Genetics when gender is female and age < 60 or gender is male and gender < 55		
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy
Random Forest (10 Trees)	53.16%	70.73%	62.28%	49.16%	92.58%	71.72%	90.60%	88.82%	89.73%
Random Forest (20 Trees)	56.24%	67.56%	62.12%	49.16%	92.58%	71.72%	90.91%	88.82%	89.89%
Random Forest (30 Trees)	54.53%	69.30%	62.20%	49.16%	92.58%	71.72%	91.22%	89.14%	90.21%
Random Forest (40 Trees)	55.90%	67.56%	61.96%	49.16%	92.58%	71.72%	90.91%	89.14%	90.05%
Random Forest (50Trees)	57.27%	66.61%	62.12%	49.16%	92.58%	71.72%	90.91%	89.14%	90.05%
C 4.5 (min number of instances/leaf: 2)	56.75%	60.92%	58.92%	52.91%	91.27%	72.84%	89.66%	96.05%	92.78%
C 4.5 (min number of instances/leaf: 5)	56.75%	60.92%	58.92%	52.91%	91.27%	72.84%	86.83%	96.38%	91.49%
C 4.5 (min number of instances/leaf: 10)	56.75%	60.92%	58.92%	52.78%	91.27%	72.78%	82.13%	96.38%	89.09%
C 4.5 (min number of instances/leaf: 15)	56.75%	60.92%	58.92%	51.88%	91.27%	72.34%	82.13%	96.38%	89.09%
C 4.5 (min number of instances/leaf: 20)	58.12%	58.39%	58.26%	50.84%	89.95%	71.16%	71.47%	93.09%	82.02%
PART (min number of instances/rule: 2)	58.63%	66.61%	62.78%	59.38%	82.89%	71.60%	90.60%	95.39%	92.94%
PART (min number of instances/rule: 5)	58.63%	66.61%	62.78%	54.33%	86.96%	71.29%	86.21%	95.72%	90.85%
PART (min number of instances/rule: 10)	58.46%	65.82%	62.28%	54.20%	87.20%	71.35%	82.13%	96.38%	89.09%
PART (min number of instances/rule: 15)	58.46%	65.82%	62.28%	51.88%	86.72%	69.98%	80.25%	96.71%	88.28%
PART (min number of instances/rule: 20)	58.46%	65.82%	62.28%	47.99%	90.43%	70.04%	75.86%	96.71%	86.04%
Decision Table (search method: Best First)	50.09%	84.02%	67.71%	57.96%	89.95%	74.58%	94.36%	95.39%	94.86%
Decision Table (search method: Greedy Stepwise)	51.28%	81.96%	67.21%	57.96%	89.95%	74.58%	94.04%	95.39%	94.70%

	Genetics when Ejection Fraction > 40			Genetics when Non Diabetic			Genetics when gender is female and age < 60 or gender is male and gender < 55		
Decision Table (search method: Linear Forward Selection)	51.79%	82.28%	67.63%	57.96%	89.95%	74.58%	94.36%	95.39%	94.86%
Decision Table (search method: Rank Search)	54.53%	81.65%	68.61%	58.09%	89.95%	74.64%	94.36%	95.39%	94.86%
Decision Table (search method: ScatterSearchV1)	61.54%	74.37%	68.20%	57.96%	89.95%	74.58%	93.10%	95.39%	94.22%
Decision Table (search method: Subset Size Forward Selection)	55.21%	77.37%	66.72%	57.70%	90.07%	74.52%	93.73%	95.39%	94.54%
Bayes Network (method for searching network structures: ICS Search Algorithm)	83.08%	50.47%	66.15%	52.39%	91.39%	72.65%	85.27%	95.72%	90.37%
Bayes Network (method for searching network structures: Naive Bayes)	71.62%	62.50%	66.89%	64.81%	75.60%	70.42%	86.21%	95.72%	90.85%
Bayes Network (method for searching network structures: Global Hill Climber)	71.62%	62.50%	66.89%	64.81%	75.60%	70.42%	86.21%	95.72%	90.85%
Bayes Network (method for searching network structures: gK2)	71.62%	62.50%	66.89%	64.81%	75.60%	70.42%	86.21%	95.72%	90.85%
Bayes Network (method for searching network structures: Global Repeated Hill Climber)	71.62%	62.50%	66.89%	64.81%	75.60%	70.42%	86.21%	95.72%	90.85%
Bayes Network (method for searching network structures: Global Tabu Search)	71.62%	62.50%	66.89%	64.81%	75.60%	70.42%	86.21%	95.72%	90.85%
Bayes Network (method for searching network structures: Local Hill Climber)	85.30%	49.21%	66.56%	64.81%	75.60%	70.42%	87.46%	95.72%	91.49%
Bayes Network (method for searching network structures: lK2)	71.62%	62.50%	66.89%	64.81%	75.60%	70.42%	86.21%	95.72%	90.85%
Bayes Network (method for searching network structures: Local LAGD Hill Climber)	85.30%	49.21%	66.56%	80.72%	53.23%	66.44%	85.89%	96.05%	90.85%
Bayes Network (method for searching network structures: Local Repeated Hill Climber)	85.30%	49.21%	66.56%	64.81%	75.60%	70.42%	87.46%	95.72%	91.49%
Bayes Network (method for searching network structures: Local Tabu Search)	85.30%	49.21%	66.56%	60.16%	79.90%	70.42%	86.21%	93.09%	89.57%
Bayes Network (method for searching network structures: Local TAN)	70.43%	63.45%	66.80%	58.86%	81.10%	70.42%	88.09%	95.72%	91.81%

Table 18: Results of several methods from datasets restricted by clinicians' feedback using SMOTE

	Diabetes Ejection Fraction AMI			Diabetes Ejection Fraction AMI Biochemical				Diabetes Ejection Fraction AMI Genetics				Diabetes Ejection Fraction AMI PUFA			
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	Improvement of accuracy	specificity	sensitivity	accuracy	Improvement of accuracy	specificity	sensitivity	accuracy	Improvement of accuracy
Bayes Network	59.22%	67.42%	63.29%	100.00%	90.82%	95.43%	32.14%	72.84%	73.81%	73.32%	10.03%	57.97%	68.95%	63.43%	0.13%
Naive Bayes	59.13%	68.95%	64.01%	69.72%	88.93%	79.27%	15.26%	69.81%	73.72%	71.75%	7.74%	57.44%	71.11%	64.23%	0.22%
Multilayer Perceptron	58.42%	72.46%	65.40%	81.75%	89.38%	85.54%	20.14%	76.05%	90.19%	83.08%	17.68%	51.56%	86.59%	68.98%	3.58%
RBF Network	60.28%	69.13%	64.68%	75.16%	86.68%	80.89%	16.20%	69.90%	70.66%	70.28%	5.60%	56.90%	70.66%	63.74%	-0.94%
K Nearest Neighbors	59.84%	69.94%	64.86%	75.33%	89.56%	82.41%	17.55%	67.14%	91.81%	79.41%	14.55%	52.72%	85.60%	69.07%	4.21%
Voting Feature Intervals	58.06%	69.49%	63.74%	79.79%	93.34%	86.53%	22.78%	54.59%	89.02%	71.71%	7.97%	56.72%	72.91%	64.77%	1.03%
Decision Table	61.71%	67.87%	64.77%	98.13%	84.61%	91.41%	26.63%	77.29%	88.21%	82.72%	17.95%	54.23%	85.42%	69.74%	4.97%
Decision Table Naive Bayes Combination	59.31%	70.39%	64.82%	99.91%	90.01%	94.99%	30.17%	77.38%	88.03%	82.68%	17.86%	54.23%	85.42%	69.74%	4.92%
RIPPER	63.22%	68.86%	66.03%	83.08%	89.74%	86.39%	20.37%	75.96%	85.33%	80.62%	14.59%	56.37%	80.47%	68.35%	2.33%
Non Nested Generalised Exemplars	63.31%	56.44%	59.89%	85.31%	80.56%	82.95%	23.05%	73.73%	80.02%	76.86%	16.97%	61.80%	56.98%	59.40%	-0.49%
PART	60.37%	69.67%	65.00%	82.01%	90.82%	86.39%	21.40%	73.82%	87.13%	80.44%	15.44%	52.27%	86.68%	69.38%	4.39%
C 4.5	60.28%	69.58%	64.91%	84.77%	88.48%	86.62%	21.71%	73.02%	87.94%	80.44%	15.53%	52.54%	86.68%	69.52%	4.61%
Random Forest	61.09%	69.22%	65.13%	87.27%	91.45%	89.35%	24.22%	74.44%	87.49%	80.93%	15.80%	53.25%	85.60%	69.34%	4.21%
Random Tree	60.64%	69.49%	65.04%	83.53%	83.17%	83.35%	18.31%	74.71%	87.76%	81.20%	16.16%	53.34%	85.60%	69.38%	4.34%

Table 19: McNemar test of several methods from datasets restricted by clinicians' feedback using SMOTE

Diabetes EF AMI

	Bayes Network	Decision Table Naive Bayes Combination	Decision Table	K Nearest Neighbors	C 4.5	RIPPER	Multilayer Perceptron	De Bayes	PART
Bayes Network	NS	S	S	S	S	S	S	S	S
Decision Table Naive Bayes Combination	S	NS	NS	NS	S	NS	S	S	S
Decision Table	S	NS	NS	NS	S	NS	S	S	S
K Nearest Neighbors	S	NS	NS	NS	S	NS	S	S	S
C 4.5	S	S	S	S	NS	S	S	NS	NS
RIPPER	S	NS	NS	NS	S	NS	S	S	S
Multilayer Perceptron	S	S	S	S	S	S	NS	S	S
Naive Bayes	S	S	S	S	NS	S	S	NS	S
PART	S	S	S	S	NS	S	S	S	NS
RBF Network	NS	S	S	S	S	S	S	S	S
Random Forest	S	S	S	S	S	NS	S	S	NS
Random Tree	S	S	S	S	S	NS	S	S	NS
Voting Feature Intervals	S	S	S	S	NS	S	S	NS	S

Diabetes EF AMI PUFA

	Decision Table Naive Bayes Combination	Decision Table	C 4.5	RIPPER	Multilayer Perceptron	PART	RBF Network	Random Forest	Random Tree
Decision Table Naive Bayes Combination	NS	NS	S	S	NS	S	S	NS	NS
Decision Table	NS	NS	S	S	NS	S	S	NS	NS
C 4.5	S	S	NS	S	NS	S	S	NS	NS
RIPPER	S	S	S	NS	S	S	S	S	S
Multilayer Perceptron	NS	NS	NS	S	NS	S	S	NS	NS
PART	S	S	S	S	S	NS	S	S	S
RBF Network	S	S	S	S	S	S	NS	S	S
Random Forest	NS	NS	NS	S	NS	S	S	NS	NS
Random Tree	NS	NS	NS	S	NS	S	S	NS	NS

Diabetes EF AMI Biochemical

	Decision Table Naive Bayes Combination	Decision Table	C 4.5	RIPPER	Multilayer Perceptron	PART	Random Forest
Decision Table Naive Bayes Combination	NS	NS	S	S	S	S	S
Decision Table	NS	NS	S	S	S	S	S
C 4.5	S	S	NS	S	NS	NS	S
RIPPER	S	S	S	NS	S	S	S
Multilayer Perceptron	S	S	NS	S	NS	NS	S
PART	S	S	NS	S	NS	NS	S
Random Forest	S	S	S	S	S	S	NS

Diabetes EF AMI Genetics

	Decision Table Naive Bayes Combination	Decision Table	C 4.5	RIPPER	Multilayer Perceptron	PART	Random Forest	Random Tree
Decision Table Naive Bayes Combination	NS	NS	NS	NS	S	NS	NS	S
Decision Table	NS	NS	NS	S	S	NS	NS	S
C 4.5	NS	NS	NS	S	S	NS	S	S
RIPPER	NS	S	S	NS	S	S	NS	NS
Multilayer Perceptron	S	S	S	S	NS	S	S	NS
PART	NS	NS	NS	S	S	NS	S	S
Random Forest	NS	NS	S	NS	S	S	NS	S
Random Tree	S	S	S	NS	NS	S	S	NS

Table 20: Results of several methods from datasets restricted by clinicians' feedback including genetics using SMOTE

	Genetics when Ejection Fraction > 40			Genetics when Non Diabetic			Genetics when gender is female and age < 60 or gender is male and gender < 55		
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy
Bayes Network	71.62%	62.50%	66.89%	64.81%	75.60%	70.42%	86.21%	95.72%	90.85%
Naive Bayes	58.29%	64.72%	61.63%	56.14%	78.35%	67.68%	88.40%	92.43%	90.37%
Multilayer Perceptron	56.07%	80.54%	68.78%	65.59%	81.82%	74.02%	95.30%	95.39%	95.35%
RBF Network	65.81%	67.88%	66.89%	62.74%	74.28%	68.74%	88.40%	93.42%	90.85%
K Nearest Neighbors	38.29%	85.92%	63.02%	46.83%	93.42%	71.04%	90.91%	93.09%	91.97%
Voting Feature Intervals	37.78%	86.23%	62.94%	38.68%	94.50%	67.68%	86.52%	95.72%	91.01%
Decision Table	50.09%	84.02%	67.71%	57.96%	89.95%	74.58%	94.36%	95.39%	94.86%
Decision Table Naive Bayes Combination	54.36%	82.59%	69.02%	58.21%	89.95%	74.70%	94.36%	95.39%	94.86%
RIPPER	45.81%	67.56%	57.11%	45.41%	92.94%	70.11%	94.67%	94.41%	94.54%
Non Nested Generalised Exemplars	61.37%	56.96%	59.08%	60.41%	68.54%	64.64%	95.92%	93.42%	94.70%
PART	53.33%	73.89%	64.01%	51.36%	91.63%	72.28%	91.22%	95.07%	93.10%
C 4.5	51.11%	66.14%	58.92%	52.91%	91.27%	72.84%	90.28%	96.05%	93.10%
Random Forest	53.16%	70.73%	62.28%	49.16%	92.58%	71.72%	90.60%	88.82%	89.73%
Random Tree	54.87%	68.99%	62.20%	50.19%	92.22%	72.03%	89.34%	89.14%	89.25%

Table 21: Methods McNemar Test of several methods from datasets restricted by clinicians' feedback including genetics using SMOTE

	Genetics when Ejection Fraction > 40							
	Bayes Network	Decision Table Naive Bayes Combination	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest	Random Tree
Bayes Network	NS	S	S	S	S	S	S	S
Decision Table Naive Bayes Combination	S	NS	NS	S	S	NS	S	S
Decision Table	S	NS	NS	S	S	NS	S	S
C 4.5	S	S	S	NS	NS	S	S	S
Multilayer Perceptron	S	S	S	NS	NS	S	S	S
PART	S	NS	NS	S	S	NS	S	S
Random Forest	S	S	S	S	S	S	NS	S
Random Tree	S	S	S	S	S	S	S	NS

	Non Diabetic Genetics				
	Bayes Network	Decision Table Naive Bayes Combination	Decision Table	Multilayer Perceptron	PART
Bayes Network	NS	S	S	S	S
Decision Table Naive Bayes Combination	S	NS	NS	NS	S
Decision Table	S	NS	NS	NS	S
Multilayer Perceptron	S	NS	NS	NS	S
PART	S	S	S	S	NS

	Genetics when gender is female and age < 60 or gender is male and gender < 55								
	Bayes Network	Decision Table Naive Bayes Combination	Decision Table	K Nearest Neighbors	C 4.5	RIPPER	Multilayer Perceptron	Non Nested Generalised Exemplars	PART
Bayes Network	NS	S	S	S	S	S	S	S	S
Decision Table Naive Bayes Combination	S	NS	NS	S	NS	NS	NS	S	NS
Decision Table	S	NS	NS	S	NS	NS	NS	S	NS
K Nearest Neighbors	S	S	S	NS	S	S	S	S	S
C 4.5	S	NS	NS	S	NS	S	S	S	NS
RIPPER	S	NS	NS	S	S	NS	NS	NS	NS
Multilayer Perceptron	S	NS	NS	S	S	NS	NS	NS	S
Non Nested Generalised Exemplars	S	S	S	S	S	NS	NS	NS	S
PART	S	NS	NS	S	NS	NS	S	S	NS

Table 22 Results of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from datasets restricted by clinicians' feedback using SMOTE

	Diabetes Ejection Fraction AMI			Diabetes Ejection Fraction AMI Biochemical				Diabetes Ejection Fraction AMI Genetics				Diabetes Ejection Fraction AMI PUFA			
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	Improvement of accuracy	specificity	sensitivity	accuracy	Improvement of accuracy	specificity	sensitivity	accuracy	Improvement of accuracy
Random Forest (10 Trees)	61.09%	69.22%	65.13%	87.27%	91.45%	89.35%	24.22%	74.44%	87.49%	80.93%	15.80%	53.25%	85.60%	69.34%	4.21%
Random Forest (20 Trees)	59.39%	69.94%	64.64%	86.55%	92.26%	89.39%	24.75%	74.35%	87.40%	80.84%	16.20%	53.25%	85.60%	69.34%	4.70%
Random Forest (30 Trees)	59.39%	69.94%	64.64%	86.55%	92.62%	89.57%	24.93%	74.71%	87.13%	80.89%	16.25%	53.25%	85.51%	69.29%	4.66%
Random Forest (40 Trees)	59.39%	69.94%	64.64%	86.73%	92.62%	89.66%	25.02%	74.62%	87.40%	80.98%	16.34%	53.25%	85.51%	69.29%	4.66%
Random Forest (50 Trees)	59.39%	69.94%	64.64%	86.38%	92.89%	89.62%	24.98%	74.80%	86.86%	80.80%	16.16%	53.25%	85.51%	69.29%	4.66%
C 4.5 (min number of instances/leaf: 2)	62.60%	67.69%	65.13%	82.28%	91.36%	86.80%	21.67%	72.40%	88.66%	80.48%	15.35%	53.43%	84.16%	68.71%	3.58%
C 4.5 (min number of instances/leaf: 5)	62.60%	67.69%	65.13%	81.75%	91.36%	86.53%	21.40%	70.26%	88.66%	79.41%	14.28%	53.43%	84.16%	68.71%	3.58%
C 4.5 (min number of instances/leaf: 10)	62.60%	67.69%	65.13%	81.48%	91.45%	86.44%	21.31%	67.94%	86.14%	76.99%	11.86%	52.81%	84.16%	68.40%	3.27%
C 4.5 (min number of instances/leaf: 15)	62.60%	67.69%	65.13%	80.85%	90.01%	85.41%	20.28%	68.30%	83.26%	75.74%	10.61%	53.43%	83.53%	68.40%	3.27%
C 4.5 (min number of instances/leaf: 20)	62.60%	67.69%	65.13%	79.25%	91.63%	85.41%	20.28%	69.46%	79.93%	74.66%	9.53%	52.89%	83.53%	68.13%	3.00%
PART (min number of instances/rule: 2)	59.13%	69.94%	64.50%	81.21%	90.46%	85.81%	21.31%	72.57%	87.13%	79.81%	15.31%	52.36%	86.50%	69.34%	4.83%
PART (min number of instances/rule: 5)	59.13%	69.94%	64.50%	82.64%	89.56%	86.08%	21.58%	71.77%	85.87%	78.78%	14.28%	52.36%	86.59%	69.38%	4.88%
PART (min number of instances/rule: 10)	59.13%	69.94%	64.50%	81.39%	91.99%	86.66%	22.16%	70.17%	83.71%	76.90%	12.40%	53.52%	84.52%	68.93%	4.43%
PART (min number of instances/rule:15)	59.13%	69.94%	64.50%	80.23%	91.18%	85.68%	21.17%	71.42%	81.55%	76.45%	11.95%	56.81%	79.57%	68.13%	3.63%
PART (min number of instances/rule: 20)	59.13%	69.94%	64.50%	80.68%	92.17%	86.39%	21.89%	67.59%	82.36%	74.93%	10.43%	56.37%	78.94%	67.59%	3.09%
Decision Table (search method: Best First)	61.71%	67.87%	64.77%	98.13%	84.61%	91.41%	26.63%	77.29%	88.21%	82.72%	17.95%	54.23%	85.42%	69.74%	4.97%
Decision Table (search method: Greedy Stepwise)	61.71%	67.87%	64.77%	98.58%	83.80%	91.23%	26.45%	77.29%	88.21%	82.72%	17.95%	55.74%	82.72%	69.16%	4.39%
Decision Table (search method: Linear Forward Selection)	61.71%	67.87%	64.77%	98.13%	84.61%	91.41%	26.63%	77.29%	88.21%	82.72%	17.95%	54.23%	85.42%	69.74%	4.97%
Decision Table (search method: Rank Search)	60.64%	68.86%	64.73%	93.59%	83.98%	88.81%	24.08%	77.29%	88.21%	82.72%	17.99%	54.23%	85.42%	69.74%	5.01%

	Diabetes Ejection Fraction AMI			Diabetes Ejection Fraction AMI Biochemical				Diabetes Ejection Fraction AMI Genetics				Diabetes Ejection Fraction AMI PUFA			
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	Improvement of accuracy	specificity	sensitivity	accuracy	Improvement of accuracy	specificity	sensitivity	accuracy	Improvement of accuracy
Decision Table (search method: ScatterSearchV1)	60.64%	68.86%	64.73%	98.93%	83.17%	91.09%	26.37%	77.29%	88.21%	82.72%	17.99%	54.23%	85.42%	69.74%	5.01%
Decision Table (search method: Subset Size Forward Selection)	61.89%	69.31%	65.58%	99.91%	82.90%	91.45%	25.87%	77.29%	88.21%	82.72%	17.14%	54.23%	85.42%	69.74%	4.16%
Bayes Network (method for searching network structures: ICS Search Algorithm)	61.89%	68.68%	65.26%	-	-	-	-	71.33%	82.00%	76.63%	11.37%	57.08%	76.87%	66.92%	1.66%
Bayes Network (method for searching network structures: Naive Bayes)	59.22%	67.42%	63.29%	100.00%	90.82%	95.43%	32.14%	72.84%	73.81%	73.32%	10.03%	57.97%	68.95%	63.43%	0.13%
Bayes Network (method for searching network structures: Global Hill Climber)	59.22%	67.42%	63.29%	100.00%	90.82%	95.43%	32.14%	72.84%	73.81%	73.32%	10.03%	57.97%	68.95%	63.43%	0.13%
Bayes Network (method for searching network structures: gK2)	59.22%	67.42%	63.29%	100.00%	90.82%	95.43%	32.14%	72.84%	73.81%	73.32%	10.03%	57.97%	68.95%	63.43%	0.13%
Bayes Network (method for searching network structures: Global Repeated Hill Climber)	59.22%	67.42%	63.29%	100.00%	90.82%	95.43%	32.14%	72.84%	73.81%	73.32%	10.03%	57.97%	68.95%	63.43%	0.13%
Bayes Network (method for searching network structures: Global Tabu Search)	59.22%	67.42%	63.29%	100.00%	90.82%	95.43%	32.14%	72.84%	73.81%	73.32%	10.03%	57.97%	68.95%	63.43%	0.13%
Bayes Network (method for searching network structures: Local Hill Climber)	59.22%	67.42%	63.29%	100.00%	90.82%	95.43%	32.14%	75.42%	67.60%	71.53%	8.24%	59.22%	67.42%	63.29%	0.00%
Bayes Network (method for searching network structures: lK2)	59.22%	67.42%	63.29%	100.00%	90.82%	95.43%	32.14%	72.84%	73.81%	73.32%	10.03%	57.97%	68.95%	63.43%	0.13%
Bayes Network (method for searching network structures: Local LAGD Hill Climber)	45.59%	79.21%	62.31%	100.00%	90.82%	95.43%	33.12%	85.04%	50.95%	68.08%	5.77%	45.59%	79.21%	62.31%	0.00%
Bayes Network (method for searching network structures: Local Repeated Hill Climber)	59.22%	67.42%	63.29%	100.00%	90.82%	95.43%	32.14%	75.42%	67.60%	71.53%	8.24%	59.22%	67.42%	63.29%	0.00%
Bayes Network (method for searching network structures: Local Tabu Search)	45.59%	79.21%	62.31%	100.00%	90.82%	95.43%	33.12%	75.42%	66.79%	71.13%	8.82%	52.63%	71.74%	62.13%	-0.18%
Bayes Network (method for searching network structures: Local TAN)	65.36%	60.67%	63.03%	100.00%	90.82%	95.43%	32.41%	72.22%	75.79%	73.99%	10.97%	58.95%	69.58%	64.23%	1.21%

Table 23: McNemar Test of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from datasets restricted by clinicians’ feedback using SMOTE

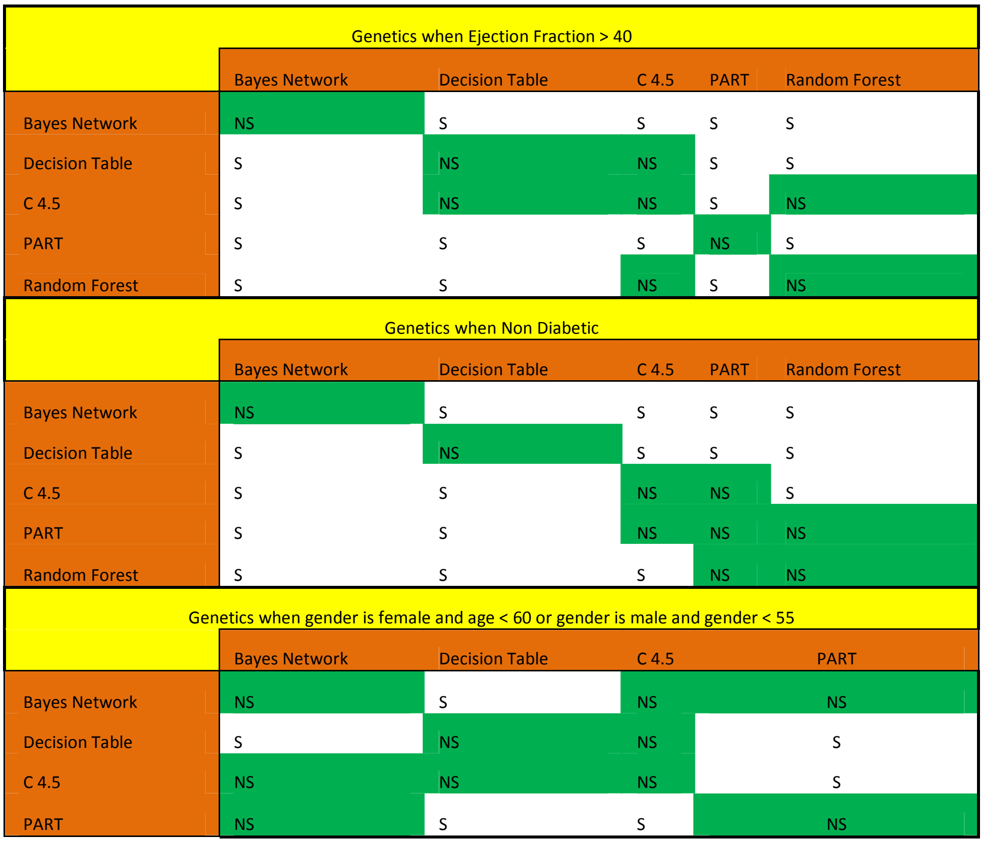
	Diabetes EF AMI						Diabetes EF AMI PUFA				
	Bayes Network	Decision Table	C 4.5	PART	Random Forest		Bayes Network	Decision Table	C 4.5	PART	Random Forest
Bayes Network	NS	S	S	S	S	Bayes Network	NS	S	S	S	S
Decision Table	S	NS	S	S	S	Decision Table	S	NS	NS	S	NS
C 4.5	S	S	NS	NS	NS	C 4.5	S	NS	NS	NS	NS
PART	S	S	NS	NS	NS	PART	S	S	NS	NS	NS
Random Forest	S	S	NS	NS	NS	Random Forest	S	NS	NS	NS	NS
	Diabetes EF AMI Biochemical						Diabetes EF AMI Genetics				
	Bayes Network	Decision Table	C 4.5	PART	Random Forest		Bayes Network	Decision Table	C 4.5	PART	Random Forest
Bayes Network	NS	NS	S	S	S	Bayes Network	NS	NS	NS	NS	NS
Decision Table	NS	NS	S	S	S	Decision Table	NS	NS	S	NS	NS
C 4.5	S	S	NS	S	S	C 4.5	NS	S	NS	S	S
PART	S	S	S	NS	S	PART	NS	NS	S	NS	S
Random Forest	S	S	S	S	NS	Random Forest	NS	NS	S	S	NS

Table 24: Results of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from datasets restricted by clinicians' feedback including genetics using SMOTE

	Genetics when Ejection Fraction > 40			Genetics when Non Diabetic			Genetics when gender is female and age < 60 or gender is male and gender < 55		
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy
Random Forest (10 Trees)	53.16%	70.73%	62.28%	49.16%	92.58%	71.72%	90.60%	88.82%	89.73%
Random Forest (20 Trees)	56.24%	67.56%	62.12%	49.16%	92.58%	71.72%	90.91%	88.82%	89.89%
Random Forest (30 Trees)	54.53%	69.30%	62.20%	49.16%	92.58%	71.72%	91.22%	89.14%	90.21%
Random Forest (40 Trees)	55.90%	67.56%	61.96%	49.16%	92.58%	71.72%	90.91%	89.14%	90.05%
Random Forest (50 Trees)	57.27%	66.61%	62.12%	49.16%	92.58%	71.72%	90.91%	89.14%	90.05%
C 4.5 (min number of instances/leaf: 2)	56.75%	60.92%	58.92%	52.91%	91.27%	72.84%	89.66%	96.05%	92.78%
C 4.5 (min number of instances/leaf: 5)	56.75%	60.92%	58.92%	52.91%	91.27%	72.84%	86.83%	96.38%	91.49%
C 4.5 (min number of instances/leaf: 10)	56.75%	60.92%	58.92%	52.78%	91.27%	72.78%	82.13%	96.38%	89.09%
C 4.5 (min number of instances/leaf: 15)	56.75%	60.92%	58.92%	51.88%	91.27%	72.34%	82.13%	96.38%	89.09%
C 4.5 (min number of instances/leaf: 20)	58.12%	58.39%	58.26%	50.84%	89.95%	71.16%	71.47%	93.09%	82.02%
PART (min number of instances/rule: 2)	58.63%	66.61%	62.78%	59.38%	82.89%	71.60%	90.60%	95.39%	92.94%
PART (min number of instances/rule: 5)	58.63%	66.61%	62.78%	54.33%	86.96%	71.29%	86.21%	95.72%	90.85%
PART (min number of instances/rule: 10)	58.46%	65.82%	62.28%	54.20%	87.20%	71.35%	82.13%	96.38%	89.09%
PART (min number of instances/rule:15)	58.46%	65.82%	62.28%	51.88%	86.72%	69.98%	80.25%	96.71%	88.28%
PART (min number of instances/rule: 20)	58.46%	65.82%	62.28%	47.99%	90.43%	70.04%	75.86%	96.71%	86.04%
Decision Table (search method: Best First)	50.09%	84.02%	67.71%	57.96%	89.95%	74.58%	94.36%	95.39%	94.86%
Decision Table (search method: Greedy Stepwise)	51.28%	81.96%	67.21%	57.96%	89.95%	74.58%	94.04%	95.39%	94.70%
Decision Table (search method: Linear Forward Selection)	51.79%	82.28%	67.63%	57.96%	89.95%	74.58%	94.36%	95.39%	94.86%

	Genetics when Ejection Fraction > 40			Genetics when Non Diabetic			Genetics when gender is female and age < 60 or gender is male and gender < 55		
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy
Decision Table (search method: Rank Search)	54.53%	81.65%	68.61%	58.09%	89.95%	74.64%	94.36%	95.39%	94.86%
Decision Table (search method: ScatterSearchV1)	61.54%	74.37%	68.20%	57.96%	89.95%	74.58%	93.10%	95.39%	94.22%
Decision Table (search method: Subset Size Forward Selection)	55.21%	77.37%	66.72%	57.70%	90.07%	74.52%	93.73%	95.39%	94.54%
Bayes Network (method for searching network structures: ICS Search Algorithm)	83.08%	50.47%	66.15%	52.39%	91.39%	72.65%	85.27%	95.72%	90.37%
Bayes Network (method for searching network structures: Naive Bayes)	71.62%	62.50%	66.89%	64.81%	75.60%	70.42%	86.21%	95.72%	90.85%
Bayes Network (method for searching network structures: Global Hill Climber)	71.62%	62.50%	66.89%	64.81%	75.60%	70.42%	86.21%	95.72%	90.85%
Bayes Network (method for searching network structures: gK2)	71.62%	62.50%	66.89%	64.81%	75.60%	70.42%	86.21%	95.72%	90.85%
Bayes Network (method for searching network structures: Global Repeated Hill Climber)	71.62%	62.50%	66.89%	64.81%	75.60%	70.42%	86.21%	95.72%	90.85%
Bayes Network (method for searching network structures: Global Tabu Search)	71.62%	62.50%	66.89%	64.81%	75.60%	70.42%	86.21%	95.72%	90.85%
Bayes Network (method for searching network structures: Local Hill Climber)	85.30%	49.21%	66.56%	64.81%	75.60%	70.42%	87.46%	95.72%	91.49%
Bayes Network (method for searching network structures: lK2)	71.62%	62.50%	66.89%	64.81%	75.60%	70.42%	86.21%	95.72%	90.85%
Bayes Network (method for searching network structures: Local LAGD Hill Climber)	85.30%	49.21%	66.56%	80.72%	53.23%	66.44%	85.89%	96.05%	90.85%
Bayes Network (method for searching network structures: Local Repeated Hill Climber)	85.30%	49.21%	66.56%	64.81%	75.60%	70.42%	87.46%	95.72%	91.49%
Bayes Network (method for searching network structures: Local Tabu Search)	85.30%	49.21%	66.56%	60.16%	79.90%	70.42%	86.21%	93.09%	89.57%
Bayes Network (method for searching network structures: Local TAN)	70.43%	63.45%	66.80%	58.86%	81.10%	70.42%	88.09%	95.72%	91.81%

Table 25: McNemar Test of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from datasets restricted by clinicians' feedback including genetics using SMOTE



Genetics when Ejection Fraction > 40

	Bayes Network	Decision Table	C 4.5	PART	Random Forest
Bayes Network	NS	S	S	S	S
Decision Table	S	NS	NS	S	S
C 4.5	S	NS	NS	S	NS
PART	S	S	S	NS	S
Random Forest	S	S	NS	S	NS

Genetics when Non Diabetic

	Bayes Network	Decision Table	C 4.5	PART	Random Forest
Bayes Network	NS	S	S	S	S
Decision Table	S	NS	S	S	S
C 4.5	S	S	NS	NS	S
PART	S	S	NS	NS	NS
Random Forest	S	S	S	NS	NS

Genetics when gender is female and age < 60 or gender is male and gender < 55

	Bayes Network	Decision Table	C 4.5	PART
Bayes Network	NS	S	NS	NS
Decision Table	S	NS	NS	S
C 4.5	NS	NS	NS	S
PART	NS	S	S	NS

After testing all the algorithms mentioned above, the results regarding the sensitivity, specificity and accuracy of the produced classifiers were provided to the clinicians, along with rules of the rule based classifiers.

The results of the classifiers produced using stratified balanced datasets were found to be accurate, but the rules produced by the rule based classifiers were not satisfying for the clinicians, mainly because of the limited number of patients in each subset.

Results produced by datasets that were balanced using the SMOTE algorithm were both accurate and the rules had clinical interpretation. After the clinicians examined the rules produced by all classifiers, once more they concluded that all rule based classifiers produced logical and non logical rules, thus the need for a classifier that could be edited remained. The classifiers that were produced by PART algorithm were preferred since they were more accurate in most datasets than the ones produced by RIPPER algorithm.

As it can be observed in Tables 18 and 22 the classifiers produced by the dataset contained Diabetes, Ejection Fraction, AMI and biochemical data are more accurate than the classifiers produced from the dataset contained Diabetes, Ejection Fraction and AMI and the dataset contained Diabetes, Ejection Fraction, AMI and genetics. The clinicians also found, after checking the rules produced by this dataset, that they were more accurate and in agreement with common medical knowledge.

In Tables 20 and 24 the results for the classifiers produced in order to predict the evolution of the disease in patients who are difficult to prognose using clinical and biochemical data are depicted; for this kind of patients, genetics data were used. From the abovementioned the biologists reviewed the rules produced by the rule based classifiers and decided that the resulted rules did not add anything to common medical knowledge. On the other hand the classifier produced by the "Diabetes, Ejection Fraction, AMI and genetics" dataset provided decision support rules that could be of help for the clinicians during the assessment of patient's condition.

In the final Decision Support System of the VPH2 platform the classifiers that will be included for the prediction of late onset heart failure will be the classifier produced using PART algorithm and the Diabetes Ejection Fraction AMI and biochemical dataset and the classifier produced using PART algorithm and the dataset that includes Diabetes, Ejection Fraction, AMI and genetics.

b. Niguarda Results

Niguarda dataset was at first split in two subsets the first subset includes patients having AMI and the second subset includes chronic patients. Three datasets for each subset were constructed; the fist dataset includes all variables except echocardiography and stress test, and it is referred as "ALL"; the second dataset referred as "ECHO" includes all variables except the stress test data; the last dataset includes all data and is referred as "STRESS TEST". In Tables 26 and 27 the variables for each dataset are shown in detail.

Niguarda dataset's target outcome is the vital status of the patient. The datasets were unbalanced, patient who finally lived were much more than those that deceased. In order to balance the datasets the SMOTE algorithm was used. The stratified balanced datasets method has not been implement yet, due to the late arrival of the dataset; stratified balanced dataset is a time consuming method. In Table 28 the results of the application of several algorithms that were applied to the AMI datasets are depicted. In Table 29 the results of the five most commonly used algorithms with several parameters' values are depicted. Similarly, in Tables 30 and 31 the results of the chronic datasets are presented. In Tables 32 and 34 the results of the algorithms are presented when the AMI datasets are balanced using the SMOTE technique. In the first table the results of the methods are presented using the default parameter values (Table 5), while in the second table the results are presented for different parameter values. In Tables 33 and 35 the corresponding McNemar tests are depicted. In the same way, the results for the chronic datasets, when they are balanced using SMOTE, are presented in Tables 36 and 38 and the corresponding McNemar tests in Tables 37 and 39.

Table 26: Variables for AMI data subset

AMI		
ALL	**ECHO**	**Stress Test**
Age	Age	Age
Sex	Sex	Sex
Smoking Habits	Body Mass Index	Body Mass Index
Hypertension	Smoking Habits	Smoking Habits
Diabetes	Hypertension	Hypertension
Dyslipidemia	Diabetes	Diabetes
Chronic kidney dysfunction	Dyslipidemia	Dyslipidemia
Dialysis	Chronic kidney dysfunction	Chronic kidney dysfunction
COPD	Dialysis	Dialysis
Atrial fibrillation history	COPD	COPD
index admissionSTENT	Atrial fibrillation history	Atrial fibrillation history
Previous STENT	index admissionSTENT	index admissionSTENT
Pre-Existing Vascular Disease	Previous STENT	Previous STENT
AMI Type	Pre-Existing Vascular Disease	Pre-Existing Vascular Disease
AMI Site	AMI Type	AMI Type
PCI	AMI Site	AMI Site
N vessels	PCI	PCI
STENT	N vessels	N vessels
CABG index admission	STENT	STENT
Number bypass	CABG index admission	CABG index admission
ACE - Inhibitors	Number bypass	Number bypass
Angiotensin-Receptor Blockers	Echocardiographic LV dilation	Echocardiographic LV dilation
Beta Blockers	LV end-Diastolic Diameter	LV end-Diastolic Diameter
Calcium Channel Blockers	LV end-Diastolic Volume	LV end-Diastolic Volume
ASA (AcetylSalicylic Acid)	LV end-Systolic Volume	LV end-Systolic Volume
Double Antiplatelet	LV Ejection Fraction	LV Ejection Fraction
Aldosterone Antag.	ACE - Inhibitors	Double product
Clopidogrel	Angiotensin-Receptor Blockers	Max Workload time
Ticlopidine	Beta Blockers	Stopping criteria
Oral anticoagulants	Calcium Channel Blockers	ACE - Inhibitors
Hypoglycaemic agents	ASA (AcetylSalicylic Acid)	Angiotensin-Receptor Blockers
Insulin	Double Antiplatelet	Beta Blockers
Statins (Lipid Lowering)	Aldosterone Antag.	Calcium Channel Blockers
Loop Diuretics	Clopidogrel	ASA (AcetylSalicylic Acid)
Digoxin	Ticlopidine	Double Antiplatelet
PUFA (ω-3)	Oral anticoagulants	Aldosterone Antag.

dose ACE/ATII inhibitors	Hypoglycaemic agents	Clopidogrel
dose Beta Blockers	Insulin	Ticlopidine
loop diuretics dose	Statins (Lipid Lowering)	Oral anticoagulants
aldosterone antagon dose	Loop Diuretics	Hypoglycaemic agents
AMI vs AMIHF	Digoxin	Insulin
Vital status (outcome to be tested)	PUFA (ω-3)	Statins (Lipid Lowering)
Date index admission	dose ACE/ATII inhibitors	Loop Diuretics
Date last follow-up	dose Beta Blockers	Digoxin
Date died	loop diuretics dose	PUFA (ω-3)
	aldosterone antagon dose	dose ACE/ATII inhibitors
	AMI vs AMIHF	dose Beta Blockers
	Vital status (outcome to be tested)	loop diuretics dose
	Date index admission	aldosterone antagon dose
	Date last follow-up	AMI vs AMIHF
	Date died	Vital status (outcome to be tested)
		Date index admission
		Date last follow-up
		Date died
Lab data		
ALT (GPT)		
aPTT		
AST (GOT)		
Blood Glucose (Serum)	**worst**	
Creatinine	**worst**	**Delta (worst-admission)**
Creatin-kinase		
Creatin-kinase MB		
Fe		
Fibrinogen	**admission**	
Gamma-GT		
Glicate Haemoglobin (blood)		
Haematocrit	**worst**	**Delta (worst-admission)**
Haemoglobin (blood)	**worst**	**Delta (worst-admission)**
HDL cholesterol	**best**	
INR		
K (K+)		
NA (NA+)		
NT Pro BNP	**worst**	

PCR		
Plateletes		
Red Blood cell counts		
Serum Total Cholesterol	**best**	
Total Bilirubine	**worst**	**Delta (worst-admission)**
Total Protein		
Triglycerides		
Troponin - T	**worst**	
Urea	**worst**	
Uric Acid	**worst**	
Ves 1h		
White Blood cell counts		

Table 27: Variables for chronic data subset

CHRONIC		
ALL	**ECHO**	**STRESS TEST**
Age	Age	Age
Sex	Sex	Sex
BMI (Body Mass Index) (calculable by Height and Weight)	BMI (Body Mass Index) (calculable by Height and Weight)	BMI (Body Mass Index) (calculable by Height and Weight)
Smoking Habits	Smoking Habits	Smoking Habits
Hypertension	Hypertension	Hypertension
Diabetes	Diabetes	Diabetes
Dyslipidemia	Dyslipidemia	Dyslipidemia
Chronic kidney dysfunction	Chronic kidney dysfunction	Chronic kidney dysfunction
Dialysis	Dialysis	Dialysis
COPD	COPD	COPD
Atrial fibrillation history	Atrial fibrillation history	Atrial fibrillation history
index admissionSTENT	index admissionSTENT	index admissionSTENT
Previous STENT	Previous STENT	Previous STENT
Pre-Existing Vascular Disease	Pre-Existing Vascular Disease	Pre-Existing Vascular Disease
Previous AMI	Previous AMI	Previous AMI
PCI	PCI	PCI
N vessels	N vessels	N vessels
STENT	STENT	STENT
CABG index admission	CABG index admission	CABG index admission
Number bypass	Number bypass	Number bypass
Mitral valve surgery	Mitral valve surgery	Mitral valve surgery
Biventricular pacing	Biventricular pacing	Biventricular pacing

Implantable Cardioverter defibrillator	Implantable Cardioverter defibrillator	Implantable Cardioverter defibrillator
Implantable Cardioverter defibrillator	Implantable Cardioverter defibrillator	Implantable Cardioverter defibrillator
BIV+ICD	BIV+ICD	BIV+ICD
ACE - Inhibitors	Echocardiographic LV dilation	Echocardiographic LV dilation
Angiotensin-Receptor Blockers	LV end-Diastolic Diameter	LV end-Diastolic Diameter
Beta Blockers	LV end-Diastolic Volume	LV end-Diastolic Volume
Calcium Channel Blockers	LV end-Systolic Volume	LV end-Systolic Volume
ASA (AcetylSalicylic Acid)	WMSI	WMSI
Double Antiplatelet	Mitral Regurgutation Severity	Mitral Regurgutation Severity
Aldosterone Antag.	LV Ejection Fraction	LV Ejection Fraction
Clopidogrel	ACE - Inhibitors	Double product
Ticlopidine	Angiotensin-Receptor Blockers	Max Workload time
Oral anticoagulants	Beta Blockers	Stopping criteria
Hypoglycaemic agents	Calcium Channel Blockers	Peak oxygen uptake (PVO2)
Insulin	ASA (AcetylSalicylic Acid)	ACE - Inhibitors
Statins (Lipid Lowering)	Double Antiplatelet	Angiotensin-Receptor Blockers
Loop Diuretics	Aldosterone Antag.	Beta Blockers
Digoxin	Clopidogrel	Calcium Channel Blockers
PUFA (ω-3)	Ticlopidine	ASA (AcetylSalicylic Acid)
dose ACE/ATII inhibitors	Oral anticoagulants	Double Antiplatelet
dose Beta Blockers	Hypoglycaemic agents	Aldosterone Antag.
loop diuretics dose	Insulin	Clopidogrel
aldosterone antagon dose	Statins (Lipid Lowering)	Ticlopidine
cIHD vs cIHF	Loop Diuretics	Oral anticoagulants
Vital status (outcome to be tested)	Digoxin	Hypoglycaemic agents
Date index admission	PUFA (ω-3)	Insulin
Date last follow-up	dose ACE/ATII inhibitors	Statins (Lipid Lowering)
Date died	dose Beta Blockers	Loop Diuretics
	loop diuretics dose	Digoxin
	aldosterone antagon dose	PUFA (ω-3)
	cIHD vs cIHF	dose ACE/ATII inhibitors
	Vital status (outcome to be tested)	dose Beta Blockers
	Date index admission	loop diuretics dose
	Date last follow-up	aldosterone antagon dose
	Date died	cIHD vs cIHF
		Vital status (outcome to be tested)
		Date index admission
		Date last follow-up
		Date died
Lab data to be appended		
Aldosterone		
ALT (GPT)		
aPTT		

AST (GOT)		
Blood Glucose (Serum)	**worst**	
Creatinine	**worst**	**Delta (worst-admission)**
Creatin-kinase		
Creatin-kinase MB		
Fe		
Fibrinogen		
Gamma-GT		
Glicate Haemoglobin (blood)		
Haematocrit	**worst**	**Delta (worst-admission)**
Haemoglobin (blood)	**worst**	**Delta (worst-admission)**
HDL cholesterol		
INR		
K (K+)	**worst**	**admission**
NA (NA+)	**worst**	**admission**
NT Pro BNP	**worst**	**Delta (discharge-worst)**
PCR		
Plateletes		
Red Blood cell counts		
Serum Total Cholesterol		
Total Bilirubine	**worst**	**Delta (worst-admission)**
Total Protein		
Triglycerides		
Troponin - T	**worst**	
Urea	**worst**	**Delta (discharge-worst)**
Uric Acid	**worst**	
Ves 1h		
White Blood cell counts		

Table 28: Results of several methods from Niguarda AMI dataset

METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy
K Nearest Neighbors	94.02%	29.76%	85.31%	95.52%	22.62%	85.63%	96.92%	20.24%	86.52%
Voting Feature Intervals	84.03%	73.81%	82.65%	85.90%	72.02%	84.02%	86.37%	72.02%	84.42%
C 4.5	97.57%	58.33%	92.25%	97.57%	58.33%	92.25%	97.57%	58.33%	92.25%
Decision Table Naive Bayes Combination	98.69%	55.36%	92.82%	98.51%	55.36%	92.66%	98.51%	55.36%	92.66%
RIPPER	98.97%	63.10%	94.11%	98.88%	61.31%	93.79%	98.23%	60.71%	93.14%
Non Nested Generalised Exemplars	98.13%	57.14%	92.57%	98.04%	56.55%	92.41%	98.04%	57.74%	92.57%
PART	96.45%	61.90%	91.77%	96.64%	60.71%	91.77%	97.20%	61.31%	92.33%
Bayes Network	86.93%	69.64%	84.58%	87.40%	69.05%	84.91%	87.21%	69.64%	84.83%
Naive Bayes	90.29%	54.76%	85.47%	89.92%	55.95%	85.31%	89.92%	55.95%	85.31%
RBF Network	97.29%	26.79%	87.73%	96.73%	26.79%	87.25%	96.55%	28.57%	87.33%
Random Tree	92.06%	52.98%	86.76%	94.21%	42.26%	87.17%	93.56%	47.62%	87.33%
Random Forest	99.07%	51.79%	92.66%	98.79%	51.79%	92.41%	99.44%	45.24%	92.09%
Decision Table	99.35%	52.98%	93.06%	99.35%	52.98%	93.06%	99.35%	52.98%	93.06%
Multilayer Perceptron	95.89%	61.31%	91.20%	96.08%	63.10%	91.61%	96.73%	66.07%	92.57%

Table 29: Results of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from Niguarda AMI dataset

	DATASET ALL			DATASET ECHO			DATASET STRESS TEST		
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy
C 4.5 (min number of instances/leaf: 2)	98.88%	56.55%	93.14%	98.88%	56.55%	93.14%	98.88%	56.55%	93.14%
C 4.5 (min number of instances/leaf: 5)	99.44%	54.76%	93.38%	99.44%	54.76%	93.38%	99.44%	54.76%	93.38%
C 4.5 (min number of instances/leaf: 10)	99.63%	55.36%	93.62%	99.63%	55.36%	93.62%	99.63%	55.36%	93.62%
C 4.5 (min number of instances/leaf: 15)	99.72%	54.17%	93.54%	99.72%	54.17%	93.54%	99.72%	54.17%	93.54%
C 4.5 (min number of instances/leaf: 20)	99.72%	54.17%	93.54%	99.72%	54.17%	93.54%	99.72%	54.17%	93.54%
PART (min number of instances/rule: 2)	98.51%	60.71%	93.38%	98.69%	61.31%	93.62%	98.69%	60.71%	93.54%
PART (min number of instances/rule: 5)	98.88%	59.52%	93.54%	98.88%	58.93%	93.46%	98.88%	58.93%	93.46%
PART (min number of instances/rule: 10)	99.72%	55.36%	93.70%	99.72%	55.36%	93.70%	99.72%	55.36%	93.70%
PART (min number of instances/rule: 15)	99.72%	55.36%	93.70%	99.72%	55.36%	93.70%	99.72%	55.36%	93.70%
PART (min number of instances/rule: 20)	99.72%	54.76%	93.62%	99.72%	54.76%	93.62%	99.72%	54.76%	93.62%
Bayes Network (method for searching network structures: lK2)	86.93%	69.64%	84.58%	87.40%	69.05%	84.91%	87.21%	69.64%	84.83%
Bayes Network (method for searching network structures: gK2)	86.93%	69.64%	84.58%	87.40%	69.05%	84.91%	87.21%	69.64%	84.83%
Bayes Network (method for searching network structures: Local TAN)	94.77%	71.43%	91.61%	94.58%	72.02%	91.53%	94.49%	70.83%	91.28%
Bayes Network (method for searching network structures: Naive Bayes)	86.93%	69.64%	84.58%	87.40%	69.05%	84.91%	87.21%	69.64%	84.83%
Bayes Network (method for searching network structures: Global Tabu Search)	92.81%	70.24%	89.75%	92.53%	69.05%	89.35%	-	-	-
Bayes Network (method for searching network structures: Local Tabu Search)	86.37%	69.64%	84.10%	86.93%	69.64%	84.58%	86.93%	69.64%	84.58%
Bayes Network (method for searching network structures: Global Hill Climber)	-	-	-	97.01%	69.05%	93.22%	-	-	-
Bayes Network (method for searching network structures: Local Hill Climber)	86.65%	69.64%	84.34%	86.65%	69.64%	84.34%	86.65%	69.64%	84.34%
Bayes Network (method for searching network structures: Local LAGD Hill Climber)	86.65%	69.64%	84.34%	86.65%	69.64%	84.34%	86.65%	69.64%	84.34%
Bayes Network (method for searching network structures: Local Repeated Hill Climber)	86.65%	69.64%	84.34%	86.65%	69.64%	84.34%	86.65%	69.64%	84.34%
Random Forest (2 Trees)	95.52%	33.33%	87.09%	97.57%	45.24%	90.48%	96.83%	35.71%	88.54%
Random Forest (10 Trees)	98.97%	51.19%	92.49%	98.88%	52.38%	92.57%	99.44%	45.24%	92.09%
Random Forest (20 Trees)	99.16%	54.76%	93.14%	99.63%	52.98%	93.30%	99.81%	50.60%	93.14%
Random Forest (30 Trees)	99.53%	55.95%	93.62%	99.63%	51.79%	93.14%	99.63%	50.60%	92.98%

	DATASET ALL			DATASET ECHO			DATASET STRESS TEST		
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy
Random Forest (40 Trees)	99.35%	56.55%	93.54%	99.91%	52.98%	93.54%	99.72%	51.19%	93.14%
Random Forest (50 Trees)	99.44%	57.14%	93.70%	99.81%	53.57%	93.54%	99.72%	51.19%	93.14%
Multilayer Perceptron (1 hidden layer neurons = [number of attributes + number of classes]/2)	95.89%	61.31%	91.20%	96.08%	63.10%	91.61%	96.73%	66.07%	92.57%
Multilayer Perceptron (1 hidden layer 2 neurons)	96.92%	59.52%	91.85%	97.01%	55.36%	91.36%	95.89%	60.12%	91.04%
Multilayer Perceptron (1 hidden layer neurons = number of attributes)	96.08%	60.71%	91.28%	96.64%	61.90%	91.93%	96.45%	62.50%	91.85%
Multilayer Perceptron (1 hidden layer neurons = number of attributes + number of classes)	96.55%	60.12%	91.61%	96.64%	58.93%	91.53%	96.36%	61.90%	91.69%
Decision Table (search method: Best First)	99.35%	52.98%	93.06%	99.35%	52.98%	93.06%	99.35%	52.98%	93.06%
Decision Table (search method: Rank Search)	99.63%	54.76%	93.54%	99.63%	54.76%	93.54%	99.63%	54.76%	93.54%
Decision Table (search method: Greedy Stepwise)	99.63%	53.57%	93.38%	99.63%	53.57%	93.38%	99.63%	53.57%	93.38%
Decision Table (search method: ScatterSearchV1)	99.63%	54.17%	93.46%	99.35%	54.76%	93.30%	99.35%	54.76%	93.30%
Decision Table (search method: Linear Forward Selection)	99.53%	54.76%	93.46%	99.35%	54.17%	93.22%	99.25%	54.17%	93.14%
Decision Table (search method: Subset Size Forward Selection)	99.25%	55.36%	93.30%	99.25%	55.36%	93.30%	99.25%	55.36%	93.30%

Table 30: Results of several methods from Niguarda chronic dataset

	DATASET ALL			DATASET ECHO			DATASET STRESS TEST		
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy
K Nearest Neighbors	89.78%	30.83%	80.63%	90.75%	24.81%	80.51%	94.06%	24.06%	83.20%
Voting Feature Intervals	81.08%	72.18%	79.70%	80.39%	72.18%	79.11%	82.04%	68.42%	79.93%
C 4.5	98.62%	54.89%	91.83%	98.62%	54.89%	91.83%	98.62%	54.89%	91.83%
Decision Table Naive Bayes Combination	99.45%	47.37%	91.37%	99.17%	47.37%	91.13%	99.17%	47.37%	91.13%
RIPPER	98.62%	50.38%	91.13%	98.62%	50.38%	91.13%	98.34%	51.88%	91.13%
Non Nested Generalised Exemplars	97.10%	55.64%	90.67%	96.69%	54.89%	90.20%	97.65%	52.63%	90.67%
PART	94.89%	57.14%	89.03%	94.75%	57.14%	88.91%	95.17%	56.39%	89.15%
Bayes Network	82.18%	73.68%	80.86%	82.32%	74.44%	81.10%	82.18%	75.19%	81.10%
Naive Bayes	88.26%	58.65%	83.66%	88.26%	60.90%	84.01%	88.12%	60.15%	83.78%
RBF Network	95.72%	28.57%	85.30%	95.72%	23.31%	84.48%	95.99%	26.32%	85.18%
Random Tree	89.78%	44.36%	82.73%	90.61%	54.89%	85.06%	92.40%	47.37%	85.41%
Random Forest	98.48%	46.62%	90.43%	98.76%	45.11%	90.43%	98.48%	44.36%	90.08%
Decision Table	99.03%	47.37%	91.02%	99.03%	47.37%	91.02%	98.76%	46.62%	90.67%
Multilayer Perceptron	93.51%	54.14%	87.40%	93.09%	51.13%	86.58%	94.75%	56.39%	88.80%

Table 31: Results of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from Niguarda chronic dataset

	DATASET ALL			DATASET ECHO			DATASET STRESS TEST		
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy
C 4.5 (min number of instances/leaf: 2)	98.90%	47.37%	90.90%	98.90%	47.37%	90.90%	98.90%	47.37%	90.90%
C 4.5 (min number of instances/leaf: 5)	99.17%	47.37%	91.13%	99.17%	47.37%	91.13%	99.17%	47.37%	91.13%
C 4.5 (min number of instances/leaf: 10)	99.86%	47.37%	91.72%	99.86%	47.37%	91.72%	99.86%	47.37%	91.72%
C 4.5 (min number of instances/leaf: 15)	99.86%	47.37%	91.72%	99.86%	47.37%	91.72%	99.86%	47.37%	91.72%
C 4.5 (min number of instances/leaf: 20)	99.86%	47.37%	91.72%	99.86%	47.37%	91.72%	99.86%	47.37%	91.72%
PART (min number of instances/rule: 2)	97.79%	48.12%	90.08%	97.38%	48.87%	89.85%	97.24%	48.87%	89.73%
PART (min number of instances/rule: 5)	98.62%	48.12%	90.78%	98.62%	48.12%	90.78%	98.62%	48.12%	90.78%
PART (min number of instances/rule: 10)	99.72%	48.87%	91.83%	99.72%	48.87%	91.83%	99.72%	48.87%	91.83%
PART (min number of instances/rule: 15)	99.72%	47.37%	91.60%	99.72%	47.37%	91.60%	99.72%	47.37%	91.60%
PART (min number of instances/rule: 20)	99.72%	47.37%	91.60%	99.72%	47.37%	91.60%	99.72%	47.37%	91.60%
Bayes Network (method for searching network structures: gK2)	82.18%	73.68%	80.86%	82.32%	74.44%	81.10%	82.18%	75.19%	81.10%
Bayes Network (method for searching network structures: IK2)	82.18%	73.68%	80.86%	82.32%	74.44%	81.10%	82.18%	75.19%	81.10%
Bayes Network (method for searching network structures: Local TAN)	92.82%	60.15%	87.75%	93.78%	60.15%	88.56%	93.78%	59.40%	88.45%
Bayes Network (method for searching network structures: Naive Bayes)	82.18%	73.68%	80.86%	82.32%	74.44%	81.10%	82.18%	75.19%	81.10%
Bayes Network (method for searching network structures: Global Tabu Search)	91.30%	66.92%	87.51%	91.44%	64.66%	87.28%	91.71%	65.41%	87.63%
Bayes Network (method for searching network structures: Local Tabu Search)	82.18%	73.68%	80.86%	82.18%	75.19%	81.10%	82.18%	75.19%	81.10%
Bayes Network (method for searching network structures: Global Hill Climber)	-	-	-	94.75%	55.64%	88.68%	-	-	-
Bayes Network (method for searching network structures: Local Hill Climber)	81.77%	74.44%	80.63%	81.49%	75.19%	80.51%	81.49%	75.19%	80.51%
Bayes Network (method for searching network structures: Local LAGD Hill Climber)	81.77%	74.44%	80.63%	81.49%	75.19%	80.51%	81.49%	75.19%	80.51%
Bayes Network (method for searching network structures: Local Repeated Hill Climber)	81.77%	74.44%	80.63%	81.49%	75.19%	80.51%	81.49%	75.19%	80.51%
Random Forest (2 Trees)	95.17%	42.86%	87.05%	95.72%	31.58%	85.76%	95.30%	36.84%	86.23%
Random Forest (10 Trees)	98.20%	46.62%	90.20%	98.62%	45.11%	90.32%	98.48%	44.36%	90.08%
Random Forest (20 Trees)	99.03%	48.12%	91.13%	99.17%	47.37%	91.13%	99.03%	45.11%	90.67%
Random Forest (30 Trees)	99.17%	50.38%	91.60%	99.31%	43.61%	90.67%	99.31%	45.11%	90.90%
Random Forest (40 Trees)	99.17%	51.13%	91.72%	99.45%	45.11%	91.02%	99.31%	45.11%	90.90%
Random Forest (50 Trees)	99.31%	51.13%	91.83%	99.31%	45.86%	91.02%	99.31%	45.11%	90.90%
Multilayer Perceptron (1 hidden layer neurons = [number of attributes + number of classes]/2)	93.51%	54.14%	87.40%	93.09%	51.13%	86.58%	94.75%	56.39%	88.80%
Multilayer Perceptron (1 hidden layer 2 neurons)	94.89%	48.87%	87.75%	95.30%	44.36%	87.40%	95.17%	51.88%	88.45%
Multilayer Perceptron (1 hidden layer neurons = number of attributes)	93.23%	54.89%	87.28%	93.65%	52.63%	87.28%	94.34%	48.12%	87.16%
Multilayer Perceptron (1 hidden layer neurons = number of attributes + number of classes)	94.34%	52.63%	87.86%	92.68%	52.63%	86.46%	94.75%	51.13%	87.98%
Decision Table (search method: Best First)	99.03%	47.37%	91.02%	99.03%	47.37%	91.02%	98.76%	46.62%	90.67%
Decision Table (search method: Rank Search)	99.72%	47.37%	91.60%	99.72%	47.37%	91.60%	99.72%	47.37%	91.60%

	DATASET ALL			DATASET ECHO			DATASET STRESS TEST		
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy
Decision Table (search method: Greedy Stepwise)	99.45%	47.37%	91.37%	99.45%	47.37%	91.37%	99.45%	47.37%	91.37%
Decision Table (search method: ScatterSearchV1)	99.45%	46.62%	91.25%	99.59%	48.12%	91.60%	99.45%	48.12%	91.48%
Decision Table (search method: Linear Forward Selection)	99.59%	47.37%	91.48%	99.45%	48.12%	91.48%	99.59%	46.62%	91.37%
Decision Table (search method: Subset Size Forward Selection)	99.86%	47.37%	91.72%	99.86%	47.37%	91.72%	99.86%	47.37%	91.72%

Table 32: Results of several methods from Niguarda AMI dataset using SMOTE.

	DATASET ALL			DATASET ECHO			DATASET STRESS TEST		
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy
K Nearest Neighbors	93.28%	89.38%	91.31%	94.86%	87.91%	91.35%	97.11%	87.64%	92.33%
Voting Feature Intervals	98.51%	91.94%	95.19%	98.51%	91.58%	95.01%	98.51%	91.67%	95.05%
C 4.5	97.85%	94.87%	96.35%	97.85%	94.87%	96.35%	97.85%	94.87%	96.35%
Decision Table Naive Bayes Combination	98.04%	93.50%	95.75%	97.95%	93.22%	95.56%	97.57%	93.96%	95.75%
RIPPER	98.51%	93.77%	96.12%	98.88%	94.23%	96.53%	98.23%	93.77%	95.98%
Non Nested Generalised Exemplars	98.23%	91.21%	94.68%	97.39%	91.12%	94.22%	98.23%	91.85%	95.01%
PART	96.17%	94.32%	95.24%	96.45%	94.32%	95.38%	95.99%	94.41%	95.19%
Bayes Network	96.73%	90.11%	93.39%	96.17%	90.48%	93.30%	96.64%	90.29%	93.44%
Naive Bayes	93.93%	89.93%	91.91%	94.21%	90.29%	92.23%	94.12%	90.48%	92.28%
RBF Network	95.33%	87.27%	91.26%	95.33%	88.46%	91.86%	94.40%	88.92%	91.63%
Random Tree	92.25%	91.48%	91.86%	89.54%	92.22%	90.89%	91.78%	91.58%	91.68%
Random Forest	98.41%	91.85%	95.10%	98.97%	91.85%	95.38%	98.69%	91.85%	95.24%
Decision Table	95.05%	89.56%	92.28%	95.61%	88.19%	91.86%	97.11%	93.13%	95.10%
Multilayer Perceptron	97.20%	93.86%	95.52%	96.73%	94.32%	95.52%	96.73%	94.23%	95.47%

Table 33: McNemar Test of several methods from Niguarda AMI dataset using SMOTE

AMI ALL METHODS	Bayes Network	Decision Table Naive Bayes Combination	Decision Table	C 4.5	RIPPER	Multilayer Perceptron	Non Nested Generalised Exemplars	PART	Random Forest
Bayes Network	NS	S	S	S	S	S	S	S	S
Decision Table Naive Bayes Combination	S	NS	S	S	NS	S	S	S	S
Decision Table	S	S	NS	S	S	S	S	S	S
C 4.5	S	S	S	NS	NS	S	S	S	S
RIPPER	S	NS	S	NS	NS	S	S	S	S
Multilayer Perceptron	S	S	S	S	S	NS	S	S	S
Non Nested Generalised Exemplars	S	S	S	S	S	S	S	S	NS
PART	S	S	S	S	S	S	S	NS	S
Random Forest	S	S	S	S	S	S	S	S	NS

AMI ECHO METHODS	Bayes Network	Decision Table Naive Bayes Combination	Decision Table	K Nearest Neighbors	C 4.5	RIPPER	Multilayer Perceptron	Non Nested Generalised Exemplars	Naïve Bayes	PART	RBF Network	Random Forest	Voting Feature Intervals
Bayes Network	NS	S	S	S	S	S	S	S	S	S	S	S	S
Decision Table Naive Bayes Combination	S	NS	S	S	NS	S	S	S	S	S	S	S	S
Decision Table	S	S	NS	S	S	S	S	S	S	S	S	S	NS
K Nearest Neighbors	S	S	S	NS	S	S	S	S	S	NS	S	S	S
C 4.5	S	NS	S	S	NS	NS	S	S	S	S	S	S	S
RIPPER	S	S	S	S	NS	NS	S	S	S	S	S	S	S
Multilayer Perceptron	S	S	S	S	S	S	NS	NS	S	S	S	S	S
Non Nested Generalised Exemplars	S	S	S	S	S	S	NS	NS	S	S	S	NS	S
Naive Bayes	S	S	S	S	S	S	S	S	NS	S	NS	S	S
	Bayes Network	Decision Table Naive Bayes Combination	Decision Table	K Nearest Neighbors	C 4.5	RIPPER	Multilayer Perceptron	Non Nested Generalised Exemplars	Naïve Bayes	PART	RBF Network	Random Forest	Voting Feature Intervals
PART	S	S	S	NS	S	S	S	S	S	NS	S	S	S
RBF Network	S	S	S	S	S	S	S	S	NS	S	NS	S	S
Random Forest	S	S	S	S	S	S	S	NS	S	S	S	NS	S
Voting Feature Intervals	S	S	NS	S	S	S	S	S	S	S	S	S	NS

AMI STRESS TEST METHODS	Bayes Network	Decision Table Naive Bayes Combination	Decision Table	C 4.5	RIPPER	Multilayer Perceptron	Non Nested Generalised Exemplars	PART	Random Forest	Voting Feature Intervals
Bayes Network	NS	S	S	S	S	S	S	S	S	S
Decision Table Naive Bayes Combination	S	NS	NS	NS	S	S	S	S	S	S
Decision Table	S	NS	NS	S	S	S	S	S	S	S
C 4.5	S	NS	S	NS	NS	S	S	S	S	S
RIPPER	S	S	S	NS	NS	S	S	S	S	S
Multilayer Perceptron	S	S	S	S	S	NS	NS	S	NS	S
Non Nested Generalised Exemplars	S	S	S	S	S	NS	NS	S	NS	S
PART	S	S	S	S	S	S	S	NS	S	S
Random Forest	S	S	S	S	S	NS	NS	S	NS	S
Voting Feature Intervals	S	S	S	S	S	S	S	S	S	NS

Table 34: Results of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from Niguarda AMI dataset using SMOTE.

	DATASET ALL			DATASET ECHO			DATASET STRESS TEST		
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy
C 4.5 (min number of instances/leaf: 2)	97.67%	94.69%	96.16%	97.57%	94.69%	96.12%	97.67%	94.69%	96.16%
C 4.5 (min number of instances/leaf: 5)	97.85%	94.60%	96.21%	97.76%	94.60%	96.16%	97.85%	94.60%	96.21%
C 4.5 (min number of instances/leaf: 10)	97.48%	94.60%	96.02%	97.76%	94.69%	96.21%	97.48%	94.60%	96.02%
C 4.5 (min number of instances/leaf: 15)	95.80%	92.77%	94.27%	95.80%	92.77%	94.27%	95.80%	92.77%	94.27%
C 4.5 (min number of instances/leaf: 20)	94.86%	92.77%	93.80%	94.86%	92.77%	93.80%	94.86%	92.77%	93.80%
PART (min number of instances/rule: 2)	98.32%	93.77%	96.02%	98.23%	93.86%	96.02%	98.32%	93.96%	96.12%
PART (min number of instances/rule: 5)	97.39%	94.14%	95.75%	97.39%	94.05%	95.70%	97.39%	94.05%	95.70%
PART (min number of instances/rule: 10)	97.67%	93.96%	95.79%	97.67%	93.96%	95.79%	97.48%	94.14%	95.79%
PART (min number of instances/rule: 15)	96.27%	94.32%	95.28%	96.27%	94.32%	95.28%	96.45%	94.41%	95.42%
PART (min number of instances/rule: 20)	94.58%	94.87%	94.73%	94.58%	94.87%	94.73%	94.40%	95.24%	94.82%
Bayes Network (method for searching network structures: gK2)	96.73%	90.11%	93.39%	96.17%	90.48%	93.30%	96.64%	90.29%	93.44%
Bayes Network (method for searching network structures: lK2)	96.73%	90.11%	93.39%	96.17%	90.48%	93.30%	96.64%	90.29%	93.44%
Bayes Network (method for searching network structures: Local TAN)	96.73%	90.11%	93.39%	96.17%	90.48%	93.30%	96.64%	90.29%	93.44%
Bayes Network (method for searching network structures: Naive Bayes)	98.32%	90.38%	94.31%	98.13%	91.03%	94.54%	97.85%	90.66%	94.22%
Bayes Network (method for searching network structures: Global Tabu Search)	96.73%	90.11%	93.39%	96.17%	90.48%	93.30%	96.64%	90.29%	93.44%
Bayes Network (method for searching network structures: Global Hill Climber)	96.73%	90.11%	93.39%	96.17%	90.48%	93.30%	96.64%	90.29%	93.44%
Bayes Network (method for searching network structures: Local Hill Climber)	99.63%	87.91%	93.71%	99.16%	88.64%	93.85%	98.97%	88.55%	93.71%
Bayes Network (method for searching network structures: Local LAGD Hill Climber)	96.64%	90.11%	93.34%	96.17%	90.48%	93.30%	96.64%	90.29%	93.44%
Bayes Network (method for searching network structures: Local Repeated Hill Climber)	96.73%	90.11%	93.39%	96.17%	90.48%	93.30%	96.64%	90.29%	93.44%
Random Forest (2 Trees)	96.27%	89.01%	92.60%	96.92%	89.29%	93.07%	95.52%	89.84%	92.65%
Random Forest (10 Trees)	98.13%	91.85%	94.96%	99.07%	91.76%	95.38%	98.69%	91.85%	95.24%
Random Forest (20 Trees)	98.69%	91.76%	95.19%	99.35%	92.22%	95.75%	99.25%	92.67%	95.93%
Random Forest (30 Trees)	99.07%	91.85%	95.42%	99.25%	92.12%	95.65%	99.53%	92.12%	95.79%
Random Forest (40 Trees)	99.25%	91.94%	95.56%	99.25%	92.03%	95.61%	99.72%	92.40%	96.02%
Random Forest (50 Trees)	99.35%	91.94%	95.61%	99.35%	92.03%	95.65%	99.53%	92.12%	95.79%
Multilayer Perceptron (1 hidden layer neurons = [number of attributes + number of classes]/2)	97.20%	93.86%	95.52%	96.73%	94.32%	95.52%	96.73%	94.23%	95.47%
Multilayer Perceptron (1 hidden layer 2 neurons)	95.70%	93.86%	94.78%	96.36%	93.77%	95.05%	95.70%	94.41%	95.05%
Multilayer Perceptron (1 hidden layer neurons = number of attributes)	97.11%	93.68%	95.38%	96.55%	94.05%	95.28%	96.55%	94.05%	95.28%
Multilayer Perceptron (1 hidden layer neurons = number of attributes + number of classes)	97.20%	94.14%	95.65%	96.83%	94.41%	95.61%	96.45%	94.78%	95.61%
Decision Table (search method: Best First)	95.05%	89.56%	92.28%	95.61%	88.19%	91.86%	97.11%	93.13%	95.10%
Decision Table (search method: Rank Search)	95.80%	89.19%	92.46%	95.80%	89.19%	92.46%	95.33%	89.93%	92.60%

	DATASET ALL			DATASET ECHO			DATASET STRESS TEST		
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy
Decision Table (search method: Greedy Stepwise)	95.05%	89.56%	92.28%	96.27%	88.10%	92.14%	97.48%	93.04%	95.24%
Decision Table (search method: ScatterSearchV1)	95.52%	94.23%	94.87%	98.60%	93.68%	96.12%	98.04%	92.95%	95.47%
Decision Table (search method: Linear Forward Selection)	95.42%	89.29%	92.33%	96.73%	89.01%	92.83%	97.57%	93.04%	95.28%
Decision Table (search method: Subset Size Forward Selection)	94.21%	88.74%	91.45%	94.30%	87.45%	90.85%	97.67%	93.04%	95.33%

Table 35: McNemar Test of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from Niguarda AMI dataset using SMOTE.

AMI ALL VALUES						
	Bayes Network	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest
Bayes Network	NS	NS	S	S	S	S
Decision Table	NS	NS	S	S	S	S
C 4.5	S	S	NS	S	NS	S
Multilayer Perceptron	S	S	S	NS	S	S
	Bayes Network	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest
PART	S	S	NS	S	NS	S
Random Forest	S	S	S	S	S	NS
AMI ECHO VALUES						
	Bayes Network	Decision Table	C 4.5	PART	Random Forest	
Bayes Network	NS	S	S	S	S	
Decision Table	S	NS	S	S	S	
C 4.5	S	S	NS	NS	S	
PART	S	S	NS	NS	S	
Random Forest	S	S	S	S	NS	
AMI STRESS TEST VALUES						
	Bayes Network	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest
Bayes Network	NS	S	S	S	S	S
Decision Table	S	NS	S	S	S	S
C 4.5	S	S	NS	S	NS	S
Multilayer Perceptron	S	S	S	NS	S	S
PART	S	S	NS	S	NS	S
Random Forest	S	S	S	S	S	NS

Table 36: Results of several methods from Niguarda chronic dataset using SMOTE.

	DATASET ALL			DATASET ECHO			DATASET STRESS TEST		
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy
K Nearest Neighbors	86.88%	88.51%	87.70%	86.74%	88.65%	87.70%	91.44%	87.69%	89.55%
Voting Feature Intervals	98.90%	76.74%	87.77%	99.17%	77.57%	88.32%	99.17%	82.22%	90.65%
C 4.5	96.55%	90.83%	93.68%	96.41%	90.83%	93.61%	96.41%	90.83%	93.61%
Decision Table Naive Bayes Combination	94.06%	92.48%	93.26%	95.17%	92.48%	93.81%	95.58%	92.48%	94.02%
RIPPER	96.96%	90.29%	93.61%	96.69%	91.11%	93.88%	97.93%	89.88%	93.88%
Non Nested Generalised Exemplars	96.69%	86.87%	91.75%	96.41%	84.40%	90.38%	97.93%	88.92%	93.40%
PART	94.61%	93.16%	93.88%	94.75%	92.89%	93.81%	95.03%	92.75%	93.88%
Bayes Network	97.10%	87.69%	92.37%	97.10%	85.77%	91.41%	97.38%	86.73%	92.03%
Naive Bayes	92.68%	88.65%	90.65%	93.51%	87.96%	90.72%	94.34%	87.96%	91.13%
RBF Network	91.57%	87.82%	89.69%	91.57%	86.87%	89.21%	91.02%	88.51%	89.76%
Random Tree	89.64%	89.33%	89.48%	91.16%	88.24%	89.69%	88.67%	89.33%	89.00%
Random Forest	96.82%	91.11%	93.95%	97.51%	90.29%	93.88%	97.65%	89.74%	93.68%
Decision Table	94.20%	88.37%	91.27%	97.65%	89.60%	93.61%	96.96%	90.42%	93.68%
Multilayer Perceptron	92.68%	91.38%	92.03%	93.37%	91.11%	92.23%	92.68%	92.34%	92.51%

Table 37: McNemar test of several methods from Niguarda chronic dataset using SMOTE

	CIHD ALL METHODS								
	Bayes Network	Decision Table Naive Bayes Combination	Decision Table	C 4.5	RIPPER	Multilayer Perceptron	Non Nested Generalised Exemplars	PART	Random Forest
Bayes Network	NS	S	S	S	S	S	S	S	S
Decision Table Naive Bayes Combination	S	NS	NS	S	S	S	S	S	S
Decision Table	S	NS	NS	S	S	S	S	S	S
C 4.5	S	S	S	NS	S	S	S	NS	S
RIPPER	S	S	S	S	NS	S	S	S	S
Multilayer Perceptron	S	S	S	S	S	NS	S	S	NS
Non Nested Generalised Exemplars	S	S	S	S	S	S	NS	S	NS
PART	S	S	S	NS	S	S	S	NS	S
Random Forest	S	S	S	S	S	NS	NS	S	NS

	CIHD ECHO METHODS							
	Bayes Network	Decision Table Naive Bayes Combination	Decision Table	C 4.5	RIPPER	Multilayer Perceptron	PART	Random Forest
Bayes Network	NS	S	S	S	S	S	S	S
Decision Table Naive Bayes Combination	S	NS	NS	S	NS	S	S	S
Decision Table	S	NS	NS	S	S	S	S	S
C 4.5	S	S	S	NS	S	S	NS	S
RIPPER	S	NS	S	S	NS	S	S	S
Multilayer Perceptron	S	S	S	S	S	NS	S	NS

PART	S	S	S	NS	S	S	NS	S
Random Forest	S	S	S	S	S	NS	S	NS

	CIHD STRESS TEST METHODS								
	Bayes Network	Decision Table Naive Bayes Combination	Decision Table	C 4.5	RIPPER	Multilayer Perceptron	Non Nested Generalised Exemplars	PART	Random Forest
Bayes Network	NS	S	S	S	S	S	S	S	S
Decision Table Naive Bayes Combination	S	NS	NS	S	NS	S	S	S	S
Decision Table	S	NS	NS	S	S	S	S	S	S
C 4.5	S	S	S	NS	S	S	S	NS	S
RIPPER	S	NS	S	S	NS	S	S	S	S
Multilayer Perceptron	S	S	S	S	S	NS	NS	S	NS
Non Nested Generalised Exemplars	S	S	S	S	S	NS	NS	S	NS
PART	S	S	S	NS	S	S	S	NS	S
Random Forest	S	S	S	S	S	NS	NS	S	NS

Table 38: Results of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from Niguarda chronic dataset using SMOTE.

	DATASET ALL			DATASET ECHO			DATASET STRESS TEST		
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy
C 4.5 (min number of instances/leaf: 2)	98.34%	90.01%	94.16%	98.34%	90.01%	94.16%	98.34%	90.01%	94.16%
C 4.5 (min number of instances/leaf: 5)	97.51%	90.29%	93.88%	97.65%	90.15%	93.88%	97.65%	90.15%	93.88%
C 4.5 (min number of instances/leaf: 10)	96.96%	90.01%	93.47%	96.96%	90.01%	93.47%	96.96%	90.01%	93.47%
C 4.5 (min number of instances/leaf: 15)	95.86%	90.56%	93.20%	95.86%	90.56%	93.20%	95.86%	90.56%	93.20%
C 4.5 (min number of instances/leaf: 20)	95.30%	89.74%	92.51%	95.03%	89.74%	92.37%	95.03%	89.74%	92.37%
PART (min number of instances/rule: 2)	96.82%	91.66%	94.23%	96.69%	91.52%	94.09%	96.27%	91.52%	93.88%
PART (min number of instances/rule: 5)	96.55%	91.38%	93.95%	96.69%	91.24%	93.95%	96.69%	91.11%	93.88%
PART (min number of instances/rule: 10)	95.72%	91.24%	93.47%	95.99%	90.97%	93.47%	96.41%	90.83%	93.61%
PART (min number of instances/rule: 15)	95.44%	91.11%	93.26%	95.44%	91.11%	93.26%	95.72%	90.97%	93.33%
PART (min number of instances/rule: 20)	95.86%	91.11%	93.47%	95.72%	90.97%	93.33%	95.58%	91.11%	93.33%
Bayes Network (method for searching network structures: gK2)	97.10%	87.69%	92.37%	97.10%	85.77%	91.41%	97.38%	86.73%	92.03%
Bayes Network (method for searching network structures: lK2)	97.10%	87.69%	92.37%	97.10%	85.77%	91.41%	97.38%	86.73%	92.03%
Bayes Network (method for searching network structures: Local TAN)	97.79%	86.46%	92.10%	97.51%	86.05%	91.75%	97.38%	86.05%	91.68%
Bayes Network (method for searching network structures: Naive Bayes)	97.10%	87.69%	92.37%	97.10%	85.77%	91.41%	97.38%	86.73%	92.03%
Bayes Network (method for searching network structures: Global Tabu Search)	96.27%	88.51%	92.37%	96.41%	88.10%	92.23%	97.38%	87.82%	92.58%
Bayes Network (method for searching network structures: Local Tabu Search)	96.55%	87.55%	92.03%	96.96%	85.91%	91.41%	97.10%	86.87%	91.96%
Bayes Network (method for searching network structures: Local Hill Climber)	96.55%	87.55%	92.03%	96.82%	86.18%	91.48%	97.10%	86.87%	91.96%
Bayes Network (method for searching network structures: Local LAGD Hill Climber)	96.55%	87.55%	92.03%	96.82%	86.18%	91.48%	97.10%	86.87%	91.96%
Bayes Network (method for searching network structures: Local Repeated Hill Climber)	96.55%	87.55%	92.03%	96.82%	86.18%	91.48%	97.10%	86.87%	91.96%
Random Forest (2 Trees)	94.75%	88.92%	91.82%	93.51%	89.19%	91.34%	94.61%	88.10%	91.34%

	DATASET ALL			DATASET ECHO			DATASET STRESS TEST		
METHOD	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy	specificity	sensitivity	accuracy
Random Forest (10 Trees)	96.69%	90.83%	93.75%	97.65%	90.42%	94.02%	97.65%	89.74%	93.68%
Random Forest (20 Trees)	97.51%	90.29%	93.88%	98.20%	89.33%	93.75%	97.93%	89.19%	93.54%
Random Forest (30 Trees)	97.38%	90.97%	94.16%	98.90%	90.29%	94.57%	98.07%	89.88%	93.95%
Random Forest (40 Trees)	98.07%	90.97%	94.50%	99.31%	89.74%	94.50%	98.20%	90.01%	94.09%
Random Forest (50 Trees)	97.93%	90.70%	94.30%	99.03%	90.15%	94.57%	98.34%	90.01%	94.16%
Multilayer Perceptron (1 hidden layer neurons = [number of attributes + number of classes]/2)	92.68%	91.38%	92.03%	93.37%	91.11%	92.23%	92.68%	92.34%	92.51%
Multilayer Perceptron (1 hidden layer 2 neurons)	92.13%	91.11%	91.62%	91.85%	92.20%	92.03%	92.27%	91.11%	91.68%
Multilayer Perceptron (1 hidden layer neurons = number of attributes)	92.40%	92.07%	92.23%	92.27%	90.97%	91.62%	93.51%	91.38%	92.44%
Multilayer Perceptron (1 hidden layer neurons = number of attributes + number of classes)	93.09%	91.11%	92.10%	92.54%	91.52%	92.03%	92.82%	91.93%	92.37%
Decision Table (search method: Best First)	94.20%	88.37%	91.27%	97.65%	89.60%	93.61%	96.96%	90.42%	93.68%
Decision Table (search method: Rank Search)	96.69%	87.82%	92.23%	96.82%	84.82%	90.79%	96.69%	85.91%	91.27%
Decision Table (search method: Greedy Stepwise)	94.20%	88.37%	91.27%	97.65%	89.60%	93.61%	98.07%	90.15%	94.09%
Decision Table (search method: ScatterSearchV1)	97.51%	90.29%	93.88%	97.38%	90.56%	93.95%	97.65%	90.56%	94.09%
Decision Table (search method: Linear Forward Selection)	94.89%	88.24%	91.55%	97.79%	89.88%	93.81%	96.96%	90.42%	93.68%
Decision Table (search method: Subset Size Forward Selection)	94.34%	88.37%	91.34%	97.93%	89.74%	93.81%	97.79%	90.42%	94.09%

Table 39 McNemar test of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from Niguarda chronic dataset using SMOTE.

	CIHD ALL VALUES					
	Bayes Network	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest
Bayes Network	NS	NS	S	S	S	S
Decision Table	NS	NS	S	S	S	S
C 4.5	S	S	NS	S	NS	S
Multilayer Perceptron	S	S	S	NS	S	S
PART	S	S	NS	S	NS	S
Random Forest	S	S	S	S	S	NS
	CIHD ECHO VALUES					
	Bayes Network	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest
Bayes Network	NS	S	S	S	S	S
Decision Table	S	NS	S	S	S	S
C 4.5	S	S	NS	S	NS	S
Multilayer Perceptron	S	S	S	NS	S	NS
PART	S	S	NS	S	NS	S
Random Forest	S	S	S	NS	S	NS
	CIHD STRESS TEST VALUES					
	Bayes Network	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest
Bayes Network	NS	S	S	S	S	S
Decision Table	S	NS	S	S	S	S
C 4.5	S	S	NS	S	NS	S
Multilayer Perceptron	S	S	S	NS	S	NS
PART	S	S	NS	S	NS	S
Random Forest	S	S	S	NS	S	NS

Results for the Niguarda datasets are accurate when the SMOTE algorithm is applied in order to balance the datasets. On unbalanced datasets the algorithm have low sensitivity, thus they do not predict patients who deceased.

Although the specificity, sensitivity and accuracy of classifiers produced from the abovementioned datasets are high, the rules are not compliant to common medical knowledge. In order to get more reasonable rules the clinicians had to further restrict the dataset. Clinicians decided to limit the datasets to one per group of patients. Moreover, the variables of each dataset were restricted too, as shown in Table 40. Furthermore, clinicians proposed to eliminate patients whose left ventricle ejection fraction was missing, since it is an important feature for the prediction. Doing so had as a result that the chronic patients' dataset was limited to 404 patients and the AMI patients' dataset was limited to 974. The datasets were still highly unbalanced, thus the SMOTE algorithm was once again applied in order to balance the datasets. The data mining algorithms previously described were applied to AMI and chronic patients' datasets.

In Tables 41and 42 the results from the application of the data mining methodologies on the first restricted version of the AMI dataset are depicted. In Table 41 the results of several methodologies using the default parameter values are shown, whereas in Table 42 the results of the methodologies using different parameter values are shown. Similarly, in Table 43 results of the application of the data mining algorithms using the default parameter values when applied to the unbalanced chronic dataset are shown and in Table 44 the results of the data mining algorithms using different parameter values are shown.

The results from the application of the data mining methodologies on the datasets balanced using SMOTE are presented in Tables 45 - 52. Tables 45 and 47 present the results of the application of the data mining methodologies using default parameter values (Table 45) and different parameter values (Table 47) of the data mining algorithms applied on the AMI dataset balanced with SMOTE. Tables 46 and 48 present the corresponding Mc Nemar tests. Likewise, Tables 49 and 51 present respectively the results of the data mining methodologies using default parameter values and different parameter values of the data mining algorithms applied on the chronic dataset balanced with SMOTE. Tables 50 and 52 present the corresponding Mc Nemar tests.

Table 40: Variables of first restricted by clinicians' version of Niguarda dataset

Variables Chronic	Variables AMI
Age	Age
Sex	Sex
BMI	BMI
Smoking Habits	Smoking Habits
Hypertension	Hypertension
Diabetes	Diabetes
Dyslipidemia	Dyslipidemia
Chronic kidney dysfunction	Chronic kidney dysfunction
Dialysis	Dialysis
COPD	COPD
Atrial fibrillation history	Atrial fibrillation history

Previous STENT			Pre-Existing Vascular Disease		
Pre-Existing Vascular Disease			AMI Type		
N vessels			AMI Site		
STENT			N vessels		
CABG index admission			STENT		
Number bypass			CABG index admission		
Biventricular pacing			Echocardiographic LV dilation		
Implantable Cardioverter defibrillator			LV Ejection Fraction		
LV end-Diastolic Volume			ACE - Inhibitors		
LV end-Systolic Volume			Angiotensin-Receptor Blockers		
LV Ejection Fraction			Beta Blockers		
ACE - Inhibitors			Calcium Channel Blockers		
Angiotensin-Receptor Blockers			ASA (AcetylSalicylic Acid)		
Beta Blockers			Double Antiplatelet		
Calcium Channel Blockers			Aldosterone Antag.		
Aldosterone Antag.			Clopidogrel		
Statins (Lipid Lowering)			Hypoglycaemic agents		
Loop Diuretics			Insulin		
loop diuretics dose			Statins (Lipid Lowering)		
cIHD vs cIHF			Loop Diuretics		
Vital status (outcome to be tested)			PUFA (ω-3)		
			AMI vs AMIHF		
			Vital status (outcome to be tested)		
Lab data			**Lab data**		
Blood Glucose (Serum)	worst		Blood Glucose (Serum)	worst	
Creatinine	worst	Delta (worst-admission)	Creatinine	worst	Delta (worst-admission)
Haemoglobin (blood)	worst	Delta (worst-admission)	Haematocrit	worst	
K (K+)	worst	admission	Haemoglobin (blood)	worst	
NA (NA+)	worst	admission	HDL cholesterol	best	
Total Bilirubine	worst		NT Pro BNP	worst	
Urea	worst		Serum Total Cholesterol	best	
Uric Acid	worst		Total Bilirubine	worst	
			Triglycerides	worst	
			Troponin - T	worst	
			Urea	worst	
			Uric Acid	worst	
			White Blood cell counts	worst	

Table 41: Results of several methods from first restricted by clinicians' version of Niguarda AMI dataset

METHOD	specificity	sensitivity	accuracy
Voting Feature Intervals	84.25%	62.60%	81.52%
RBF Network	97.77%	17.07%	87.58%
Random Tree	90.60%	25.20%	82.34%
Random Forest	98.59%	13.01%	87.78%
PART	93.18%	32.52%	85.52%
Non Nested Generalised Exemplars	95.65%	23.58%	86.55%
Naive Bayes	86.13%	52.03%	81.83%
Multilayer Perceptron	94.24%	37.40%	87.06%
RIPPER	95.53%	21.14%	86.14%
C 4.5	94.95%	26.83%	86.35%
K Nearest Neighbors	96.71%	22.76%	87.37%
Decision Table Naive Bayes Combination	94.59%	15.45%	84.60%
Decision Table	98.35%	7.32%	86.86%
Bayes Network	85.55%	61.79%	82.55%

Table 42: Results of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from first restricted by clinicians' version of Niguarda AMI dataset

METHOD	specificity	sensitivity	accuracy
Bayes Network (method for searching network structures: Global Hill Climber)	91.42%	47.15%	85.83%
Bayes Network (method for searching network structures: gK2)	85.55%	61.79%	82.55%
Bayes Network (method for searching network structures: ICS Search Algorithm)	90.13%	31.71%	82.75%
Bayes Network (method for searching network structures: Local Hill Climber)	86.25%	64.23%	83.47%
Bayes Network (method for searching network structures: lK2)	85.55%	61.79%	82.55%
Bayes Network (method for searching network structures: Local LAGD Hill Climber)	86.25%	64.23%	83.47%
Bayes Network (method for searching network structures: Local Repeated Hill Climber)	86.25%	64.23%	83.47%
Bayes Network (method for searching network structures: Local Tabu Search)	85.43%	61.79%	82.44%
Bayes Network (method for searching network structures: Local TAN)	91.77%	45.53%	85.93%
Bayes Network (method for searching network structures: Naive Bayes)	85.55%	61.79%	82.55%
C 4.5 (min number of instances/leaf: 10)	98.82%	2.44%	86.65%
C 4.5 (min number of instances/leaf: 15)	98.00%	4.07%	86.14%
C 4.5 (min number of instances/leaf: 2)	97.18%	9.76%	86.14%
C 4.5 (min number of instances/leaf: 20)	98.47%	1.63%	86.24%
C 4.5 (min number of instances/leaf: 5)	98.24%	4.07%	86.35%
Decision Table (search method: Best First)	98.35%	7.32%	86.86%
Decision Table (search method: Greedy Stepwise)	98.35%	6.50%	86.76%
Decision Table (search method: Linear Forward Selection)	98.71%	4.88%	86.86%

METHOD	specificity	sensitivity	accuracy
Decision Table (search method: Rank Search)	97.41%	11.38%	86.55%
Decision Table (search method: ScatterSearchV1)	99.29%	5.69%	87.47%
Decision Table (search method: Subset Size Forward Selection)	99.18%	6.50%	87.47%
Multilayer Perceptron (1 hidden layer neurons = [number of attributes + number of classes]/2)	94.24%	37.40%	87.06%
Multilayer Perceptron (1 hidden layer neurons = number of attributes + number of classes)	93.65%	38.21%	86.65%
Multilayer Perceptron (1 hidden layer neurons = number of attributes)	94.36%	39.02%	87.37%
Multilayer Perceptron (1 hidden layer 2 neurons)	94.24%	31.71%	86.35%
PART (min number of instances/rule: 10)	97.65%	8.94%	86.45%
PART (min number of instances/rule: 15)	97.88%	8.13%	86.55%
PART (min number of instances/rule: 2)	95.42%	16.26%	85.42%
PART (min number of instances/rule: 20)	97.65%	10.57%	86.65%
PART (min number of instances/rule: 5)	97.77%	8.13%	86.45%
Random Forest (10 Trees)	98.59%	13.01%	87.78%
Random Forest (2 Trees)	95.89%	16.26%	85.83%
Random Forest (20 Trees)	98.94%	12.20%	87.99%
Random Forest (30 Trees)	98.59%	13.82%	87.89%
Random Forest (40 Trees)	98.71%	11.38%	87.68%
Random Forest (50 Trees)	98.71%	11.38%	87.68%

Table 43: Results of several methods from first restricted by clinicians' version of Niguarda chronic dataset.

METHOD	specificity	sensitivity	accuracy
Bayes Network	83.28%	52.17%	77.97%
C 4.5	94.03%	21.74%	81.68%
Decision Table	97.31%	7.25%	81.93%
Decision Table Naive Bayes Combination	89.55%	13.04%	76.49%
K Nearest Neighbors	85.97%	24.64%	75.50%
Multilayer Perceptron	89.55%	27.54%	78.96%
Naive Bayes	82.39%	55.07%	77.72%
Non Nested Generalised Exemplars	94.33%	13.04%	80.45%
PART	85.67%	30.43%	76.24%
Random Forest	95.82%	24.64%	83.66%
Random Tree	85.97%	27.54%	75.99%
RBF Network	91.64%	27.54%	80.69%
RIPPER	92.24%	27.54%	81.19%
Voting Feature Intervals	74.63%	60.87%	72.28%

Table 44: Results of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from first restricted by clinicians' version of Niguarda chronic dataset

METHOD	specificity	sensitivity	accuracy
Bayes Network (method for searching network structures: Global Tabu Search)	89.55%	31.88%	79.70%
Bayes Network (method for searching network structures: Global Hill Climber)	89.25%	33.33%	79.70%
Bayes Network (method for searching network structures: gK2)	83.28%	52.17%	77.97%
Bayes Network (method for searching network structures: Global Repeated Hill Climber)	89.25%	33.33%	79.70%
Bayes Network (method for searching network structures: ICS Search Algorithm)	86.57%	37.68%	78.22%
Bayes Network (method for searching network structures: Local Hill Climber)	80.60%	53.62%	75.99%
Bayes Network (method for searching network structures: lK2)	83.28%	52.17%	77.97%
Bayes Network (method for searching network structures: Local LAGD Hill Climber)	80.90%	53.62%	76.24%
Bayes Network (method for searching network structures: Local Repeated Hill Climber)	80.60%	53.62%	75.99%
Bayes Network (method for searching network structures: Local Tabu Search)	81.19%	56.52%	76.98%
Bayes Network (method for searching network structures: Local TAN)	89.85%	40.58%	81.44%
Bayes Network (method for searching network structures: Naive Bayes)	83.28%	52.17%	77.97%
C 4.5 (min number of instances/leaf: 10)	99.40%	1.45%	82.67%
C 4.5 (min number of instances/leaf: 15)	99.40%	0.00%	82.43%
C 4.5 (min number of instances/leaf: 2)	98.51%	8.70%	83.17%
C 4.5 (min number of instances/leaf: 20)	100.00%	0.00%	82.92%
C 4.5 (min number of instances/leaf: 5)	99.40%	7.25%	83.66%
Decision Table (search method: Best First)	97.31%	7.25%	81.93%
Decision Table (search method: Greedy Stepwise)	99.40%	2.90%	82.92%
Decision Table (search method: Linear Forward Selection)	96.12%	4.35%	80.45%
Decision Table (search method: Rank Search)	96.72%	15.94%	82.92%
Decision Table (search method: ScatterSearchV1)	99.40%	2.90%	82.92%
Decision Table (search method: Subset Size Forward Selection)	99.70%	2.90%	83.17%
Multilayer Perceptron (1 hidden layer neurons = [number of attributes + number of classes]/2)	89.55%	27.54%	78.96%
Multilayer Perceptron (1 hidden layer neurons = number of attributes + number of classes)	90.75%	33.33%	80.94%
Multilayer Perceptron (1 hidden layer neurons = number of attributes)	89.25%	36.23%	80.20%
Multilayer Perceptron (1 hidden layer 2 neurons)	91.34%	28.99%	80.69%
PART (min number of instances/rule: 10)	96.12%	7.25%	80.94%
PART (min number of instances/rule: 15)	99.10%	0.00%	82.18%
PART (min number of instances/rule: 2)	94.33%	15.94%	80.94%
PART (min number of instances/rule: 20)	99.40%	0.00%	82.43%
PART (min number of instances/rule: 5)	95.82%	20.29%	82.92%
Random Forest (10 Trees)	95.82%	24.64%	83.66%
Random Forest (2 Trees)	91.64%	15.94%	78.71%
Random Forest (20 Trees)	97.91%	18.84%	84.41%
Random Forest (30 Trees)	97.61%	14.49%	83.42%
Random Forest (40 Trees)	98.51%	17.39%	84.65%
Random Forest (50 Trees)	97.91%	14.49%	83.66%

Table 45: Results of several methods from first restricted by clinicians' version of Niguarda AMI dataset using SMOTE.

METHOD	specificity	sensitivity	accuracy
Bayes Network	93.30%	90.46%	91.88%
C 4.5	94.71%	89.05%	91.88%
Decision Table	96.71%	89.05%	92.88%
Decision Table Naive Bayes Combination	95.06%	88.93%	92.00%
K Nearest Neighbors	96.12%	87.51%	91.82%
Multilayer Perceptron	91.89%	91.64%	91.76%
Naive Bayes	94.01%	89.63%	91.82%
Non Nested Generalised Exemplars	96.12%	86.10%	91.12%

METHOD	specificity	sensitivity	accuracy
PART	94.83%	90.69%	92.76%
Random Forest	97.41%	88.46%	92.94%
Random Tree	89.07%	89.99%	89.53%
RBF Network	97.30%	87.87%	92.59%
RIPPER	98.59%	87.63%	93.12%
Voting Feature Intervals	98.82%	86.45%	92.65%

Table 46: McNemar test of several methods from first restricted by clinicians' version of Niguarda AMI dataset using SMOTE.

	Decision Table Naive Bayes Combination	Decision Table	K Nearest Neighbors	C 4.5	RIPPER	Multilayer Perceptron	Non Nested Generalised Exemplars	Naïve Bayes	PART	RBF Network	Random Forest	Voting Feature Intervals	Bayes Network
Decision Table Naive Bayes Combination	NS	NS	S	S	NS	S	S	NS	S	NS	S	NS	NS
Decision Table	NS	NS	S	S	NS	S	S	S	S	NS	S	NS	S
K Nearest Neighbors	S	S	NS	S	S	NS	NS	S	S	S	NS	S	S
C 4.5	S	S	S	NS	S	S	S	S	S	S	S	S	S
RIPPER	NS	NS	S	S	NS	S	S	S	S	NS	S	NS	S
Multilayer Perceptron	S	S	NS	S	S	NS	NS	S	S	S	NS	S	S
Non Nested Generalised Exemplars	S	S	NS	S	S	NS	NS	S	S	S	NS	S	S
Naïve Bayes	NS	S	S	S	S	S	S	NS	S	NS	S	NS	NS
PART	S	S	S	S	S	S	S	S	NS	S	S	S	S
RBF Network	NS	NS	S	S	NS	S	S	NS	S	NS	S	NS	NS
Random Forest	S	S	NS	S	S	NS	NS	S	S	S	NS	S	S
Voting Feature Intervals	NS	NS	S	S	NS	S	S	NS	S	NS	S	NS	NS
Bayes Network	NS	S	S	S	S	S	S	NS	S	NS	S	NS	NS

Table 47: Results of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from first restricted by clinicians' version of Niguarda AMI dataset using SMOTE

METHOD	specificity	sensitivity	accuracy
Bayes Network (method for searching network structures: Global Hill Climber)	94.83%	90.46%	92.65%
Bayes Network (method for searching network structures: gK2)	93.30%	90.46%	91.88%
Bayes Network (method for searching network structures: Local Hill Climber)	92.95%	90.46%	91.71%
Bayes Network (method for searching network structures: lK2)	93.30%	90.46%	91.88%
Bayes Network (method for searching network structures: Local LAGD Hill Climber)	93.07%	90.46%	91.76%

METHOD	specificity	sensitivity	accuracy
Bayes Network (method for searching network structures: Local Repeated Hill Climber)	92.95%	90.46%	91.71%
Bayes Network (method for searching network structures: Local Tabu Search)	92.95%	90.46%	91.71%
Bayes Network (method for searching network structures: Local TAN)	97.65%	87.04%	92.35%
Bayes Network (method for searching network structures: Naive Bayes)	93.30%	90.46%	91.88%
C 4.5 (min number of instances/leaf: 10)	96.00%	88.10%	92.06%
C 4.5 (min number of instances/leaf: 15)	95.42%	88.81%	92.12%
C 4.5 (min number of instances/leaf: 2)	96.94%	87.87%	92.41%
C 4.5 (min number of instances/leaf: 20)	95.77%	88.46%	92.12%
C 4.5 (min number of instances/leaf: 5)	96.24%	88.93%	92.59%
Decision Table (search method: Best First)	96.71%	89.05%	92.88%
Decision Table (search method: Greedy Stepwise)	96.94%	88.57%	92.76%
Decision Table (search method: Linear Forward Selection)	96.71%	89.05%	92.88%
Decision Table (search method: Rank Search)	97.06%	87.40%	92.24%
Decision Table (search method: ScatterSearchV1)	96.71%	87.87%	92.29%
Decision Table (search method: Subset Size Forward Selection)	97.77%	88.34%	93.06%
Multilayer Perceptron (1 hidden layer neurons = [number of attributes + number of classes]/2)	91.89%	91.64%	91.76%
Multilayer Perceptron (1 hidden layer neurons = number of attributes + number of classes)	93.54%	91.05%	92.29%
Multilayer Perceptron (1 hidden layer neurons = number of attributes)	92.48%	91.05%	91.76%
Multilayer Perceptron (1 hidden layer 2 neurons)	94.71%	90.81%	92.76%
PART (min number of instances/rule: 10)	95.65%	88.81%	92.24%
PART (min number of instances/rule: 15)	95.89%	88.10%	92.00%
PART (min number of instances/rule: 2)	95.06%	89.05%	92.06%
PART (min number of instances/rule: 20)	96.00%	88.34%	92.18%
PART (min number of instances/rule: 5)	96.12%	88.46%	92.29%
Random Forest (10 Trees)	97.41%	88.46%	92.94%
Random Forest (2 Trees)	96.59%	87.16%	91.88%
Random Forest (20 Trees)	97.88%	88.46%	93.18%
Random Forest (30 Trees)	98.00%	88.34%	93.18%
Random Forest (40 Trees)	98.12%	88.10%	93.12%
Random Forest (50 Trees)	98.24%	88.22%	93.24%

Table 48: McNemar test of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from first restricted by clinicians' version of Niguarda AMI dataset using SMOTE

	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest	Bayes Network
Decision Table	NS	NS	S	NS	S	NS
C 4.5	NS	NS	S	NS	S	NS
Multilayer Perceptron	S	S	S	S	S	S
PART	NS	NS	S	NS	S	NS
Random Forest	S	S	S	S	NS	S
Bayes Network	NS	NS	S	NS	S	NS

Table 49: Results of several methods from first restricted by clinicians' version of Niguarda chronic dataset using SMOTE.

METHOD	specificity	sensitivity	accuracy
Bayes Network	91.94%	86.96%	89.26%
C 4.5	93.13%	84.65%	88.57%
Decision Table	90.75%	85.68%	88.02%
Decision Table Naive Bayes Combination	89.55%	87.98%	88.71%
K Nearest Neighbors	80.90%	82.61%	81.82%
Multilayer Perceptron	87.46%	88.24%	87.88%
Naive Bayes	85.37%	89.51%	87.60%
Non Nested Generalised Exemplars	95.52%	63.17%	78.10%
PART	87.76%	87.47%	87.60%
Random Forest	95.22%	85.42%	89.94%
Random Tree	86.87%	87.47%	87.19%
RBF Network	91.94%	85.68%	88.57%
RIPPER	94.63%	84.65%	89.26%
Voting Feature Intervals	98.51%	77.49%	87.19%

Table 50: McNemar test of several methods from first restricted by clinicians' version of Niguarda chronic dataset using SMOTE.

	Decision Table Naive Bayes Combination	Decision Table	C 4.5	RIPPER	Multilayer Perceptron	Naive Bayes	PART	RBF Network	Random Forest	Random Tree	Voting Feature Intervals	Bayes Network
Decision Table Naive Bayes Combination	NS	NS	NS	NS	S	S	S	NS	S	S	NS	NS
Decision Table	NS	NS	S	NS	S	S	S	NS	S	S	NS	NS
C 4.5	NS	S	NS	NS	S	S	S	S	S	S	S	S
RIPPER	NS	NS	NS	NS	S	S	S	NS	S	S	S	NS
Multilayer Perceptron	S	S	S	S	NS	S	S	S	NS	NS	S	S
Naive Bayes	S	S	S	S	S	NS	S	S	S	S	NS	S
PART	S	S	S	S	S	S	NS	S	S	S	S	S
RBF Network	NS	NS	S	NS	S	S	S	NS	S	S	NS	NS
Random Forest	S	S	S	S	NS	S	S	S	NS	NS	S	S
Random Tree	S	S	S	S	NS	S	S	S	NS	NS	S	S
Voting Feature Intervals	NS	NS	S	S	S	NS	S	NS	S	S	NS	NS
Bayes Network	NS	NS	S	NS	S	S	S	NS	S	S	NS	NS

Table 51: : Results of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from first restricted by clinicians' version of Niguarda chronic dataset using SMOTE

METHOD	specificity	sensitivity	accuracy
Bayes Network (method for searching network structures: Global Tabu Search)	92.24%	87.72%	89.81%
Bayes Network (method for searching network structures: Global Hill Climber)	93.13%	87.72%	90.22%
Bayes Network (method for searching network structures: gK2)	91.94%	86.96%	89.26%
Bayes Network (method for searching network structures: Global Repeated Hill Climber)	93.13%	87.72%	90.22%
Bayes Network (method for searching network structures: Local Hill Climber)	91.94%	86.96%	89.26%
Bayes Network (method for searching network structures: lK2)	91.94%	86.96%	89.26%
Bayes Network (method for searching network structures: Local LAGD Hill Climber)	91.94%	86.96%	89.26%
Bayes Network (method for searching network structures: Local Repeated Hill Climber)	91.94%	86.96%	89.26%
Bayes Network (method for searching network structures: Local Tabu Search)	91.94%	86.96%	89.26%
Bayes Network (method for searching network structures: Local TAN)	95.82%	85.42%	90.22%
Bayes Network (method for searching network structures: Bayes)	91.94%	86.96%	89.26%
C 4.5 (min number of instances/leaf: 10)	91.34%	85.68%	88.29%
C 4.5 (min number of instances/leaf: 15)	90.75%	85.93%	88.15%
C 4.5 (min number of instances/leaf: 2)	92.84%	85.68%	88.98%
C 4.5 (min number of instances/leaf: 20)	90.45%	86.19%	88.15%
C 4.5 (min number of instances/leaf: 5)	91.94%	85.68%	88.57%
Decision Table (search method: Best First)	90.75%	85.68%	88.02%
Decision Table (search method: Greedy Stepwise)	90.75%	85.68%	88.02%
Decision Table (search method: Linear Forward Selection)	91.64%	85.93%	88.57%
Decision Table (search method: Rank Search)	92.24%	86.45%	89.12%
Decision Table (search method: ScatterSearchV1)	96.12%	84.14%	89.67%
Decision Table (search method: Subset Size Forward Selection)	93.43%	85.93%	89.39%
Multilayer Perceptron (1 hidden layer neurons = [number of attributes + number of classes]/2)	87.46%	88.24%	87.88%
Multilayer Perceptron (1 hidden layer neurons = number of attributes + number of classes)	88.06%	88.24%	88.15%
Multilayer Perceptron (1 hidden layer neurons = number of attributes)	86.27%	88.75%	87.60%
Multilayer Perceptron (1 hidden layer 2 neurons)	88.36%	88.49%	88.43%
PART (min number of instances/rule: 10)	92.24%	85.93%	88.84%
PART (min number of instances/rule: 15)	89.55%	86.45%	87.88%
PART (min number of instances/rule: 2)	91.94%	85.68%	88.57%
PART (min number of instances/rule: 20)	89.55%	86.70%	88.02%
PART (min number of instances/rule: 5)	91.04%	87.47%	89.12%
Random Forest (10 Trees)	95.22%	85.42%	89.94%
Random Forest (2 Trees)	91.04%	84.14%	87.33%
Random Forest (20 Trees)	95.52%	85.93%	90.36%
Random Forest (30 Trees)	96.12%	85.93%	90.63%
Random Forest (40 Trees)	96.12%	85.68%	90.50%
Random Forest (50 Trees)	96.42%	85.42%	90.50%

Table 52: McNemar test of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from first restricted by clinicians' version of Niguarda chronic dataset using SMOTE.

	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest	Bayes Network
Decision Table	NS	NS	S	NS	S	NS
C 4.5	NS	NS	S	NS	S	NS
Multilayer Perceptron	S	S	NS	S	S	S
PART	NS	NS	S	NS	S	NS
Random Forest	S	S	S	S	NS	S
Bayes Network	NS	NS	S	NS	S	NS

The results from the application of the data mining methodologies were presented to the clinicians along with the rules produced from the rule based classifiers. The application of the data mining methodologies on the unbalanced datasets was poor in sensitivity, thus classifiers did not predict correctly patients who deceased, both in AMI and chronic dataset.

Results from the data mining methodologies applied on AMI and chronic datasets balanced using SMOTE algorithm were more accurate and had larger sensitivity. Clinicians reviewed the rules produced by PART algorithm and decided that most of them stand to the common sense-common knowledge test, although rules produced from the AMI dataset were too “broad” to be useful in the extraction of new knowledge.

After reviewing the above mentioned results the clinicians proposed to check the accuracy of the classifiers when the drug treatment of the patients is not included. Two new datasets were constructed using the features shown in Table 53. In the next pages the results from the application of the data mining methodologies on the second version of the restricted AMI and chronic dataset are presented. In Table 54 the results of the methodologies using the default parameters values on the unbalanced AMI dataset are presented and Table 55 different parameter values are tested in order to find the one giving best result. Similarly, in Table 56 the results from the application of the algorithms using default parameter values on the unbalanced chronic dataset are presented, while in Table 57 the results when using different parameter values are presented. In each table the last column referred as improvement shows the difference in accuracy between the current dataset and the first restricted version.

The next step was to balance the datasets using SMOTE algorithm. In Table 58 and Table 60 the results of the application of the data mining algorithms using default parameter values and several parameter values respectively on the AMI dataset balanced with SMOTE are depicted. Tables 59 and 61 present the corresponding McNemar tests, for the abovementioned methodologies. Tables 62 - 65 depict the results of the data mining methodologies on the chronic dataset balanced with SMOTE. Table 62 shows the results of the methodologies when using the default parameter values and Table 63 the corresponding McNemar test.

Table 64 shows the results of the methodologies when different parameter values are applied and Table 65 the corresponding McNemar test.

Table 53: Variables of second restricted by clinicians’ version of Niguarda dataset

Variables Chronic	Variables AMI
Age	Age
Sex	Sex
BMI	BMI
Smoking Habits	Smoking Habits
Hypertension	Hypertension
Diabetes	Diabetes
Dyslipidemia	Dyslipidemia
Chronic kidney dysfunction	Chronic kidney dysfunction
Dialysis	Dialysis
COPD	COPD
Atrial fibrillation history	Atrial fibrillation history
Previous STENT	Pre-Existing Vascular Disease

Pre-Existing Vascular Disease			AMI Type		
N vessels			AMI Site		
STENT			N vessels		
CABG index admission			STENT		
Number bypass			CABG index admission		
Biventricular pacing			Echocardiographic LV dilation		
Implantable Cardioverter defibrillator			LV Ejection Fraction		
LV end-Diastolic Volume			AMI vs AMIHF		
LV end-Systolic Volume			Vital status (outcome to be tested)		
LV Ejection Fraction					
cIHD vs cIHF					
Vital status (outcome to be tested)					
Lab data			**Lab data**		
Blood Glucose (Serum)	worst		Blood Glucose (Serum)	worst	
Creatinine	worst	Delta (worst-admission)	Creatinine	worst	Delta (worst-admission)
Haemoglobin (blood)	worst	Delta (worst-admission)	Haematocrit	worst	
K (K+)	worst	admission	Haemoglobin (blood)	worst	
NA (NA+)	worst	admission	HDL cholesterol	best	
Total Bilirubine	worst		NT Pro BNP	worst	
Urea	worst		Serum Total Cholesterol	best	
Uric Acid	worst		Total Bilirubine	worst	
			Triglycerides	worst	
			Troponin - T	worst	
			Urea	worst	
			Uric Acid	worst	
			White Blood cell counts	worst	

Table 54: Results of several methods from second restricted by clinicians' version of Niguarda AMI dataset

METHOD	specificity	sensitivity	accuracy	IMPROVEMENT
K Nearest Neighbors	94.10%	22.13%	85.05%	-2.32%
Voting Feature Intervals	85.50%	57.38%	81.96%	0.44%
C 4.5	95.75%	25.41%	86.91%	0.56%
Decision Table Naive Bayes Combination	95.05%	12.30%	84.64%	0.04%
RIPPER	96.93%	12.30%	86.29%	0.15%
Non Nested Generalised Exemplars	96.58%	22.13%	87.22%	0.67%
PART	94.22%	27.87%	85.88%	0.35%
Bayes Network	87.26%	54.92%	83.20%	0.65%
Naive Bayes	84.79%	55.74%	81.13%	-0.69%
RBF Network	95.99%	19.67%	86.39%	-1.19%

METHOD	specificity	sensitivity	accuracy	IMPROVEMENT
Random Tree	91.86%	24.59%	83.40%	1.06%
Random Forest	97.88%	14.75%	87.42%	-0.36%
Decision Table	98.58%	6.56%	87.01%	0.15%
Multilayer Perceptron	93.75%	36.07%	86.49%	-0.57%

Table 55: Results of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from second restricted by clinicians' version of Niguarda AMI dataset

METHOD	specificity	sensitivity	accuracy	IMPROVEMENT
Bayes Network (method for searching network structures: CI Search Algorithm	87.26%	54.92%	83.20%	
Bayes Network (method for searching network structures: Global Tabu Search	91.75%	42.62%	85.57%	85.57%
Bayes Network (method for searching network structures: Global Hill Climber	91.75%	42.62%	85.57%	-0.26%
Bayes Network (method for searching network structures: gK2	87.26%	54.92%	83.20%	0.65%
Bayes Network (method for searching network structures: Global Repeated Hill Climber	91.75%	42.62%	85.57%	85.57%
Bayes Network (method for searching network structures: ICS Search Algorithm	90.09%	35.25%	83.20%	0.44%
Bayes Network (method for searching network structures: Local Hill Climber	86.91%	54.92%	82.89%	-0.58%
Bayes Network (method for searching network structures: lK2	87.26%	54.92%	83.20%	0.65%
Bayes Network (method for searching network structures: Local LAGD Hill Climber	86.91%	54.92%	82.89%	-0.58%
Bayes Network (method for searching network structures: Local Repeated Hill Climber	86.91%	54.92%	82.89%	-0.58%
Bayes Network (method for searching network structures: Local Simulated Annealing	94.69%	23.77%	85.77%	85.77%
Bayes Network (method for searching network structures: Local Tabu Search	86.91%	54.92%	82.89%	0.44%
Bayes Network (method for searching network structures: Local TAN	92.69%	44.26%	86.60%	0.66%
Bayes Network (method for searching network structures: Naive Bayes	87.26%	54.92%	83.20%	0.65%
C 4.5 (min number of instances/leaf: 10)	98.58%	6.56%	87.01%	0.36%
C 4.5 (min number of instances/leaf: 15)	98.47%	5.74%	86.80%	0.66%
C 4.5 (min number of instances/leaf: 2)	98.70%	4.92%	86.91%	0.77%
C 4.5 (min number of instances/leaf: 20)	98.82%	4.10%	86.91%	0.66%
C 4.5 (min number of instances/leaf: 5)	97.52%	10.66%	86.60%	0.25%
Decision Table (search method: Best First	99.41%	6.56%	87.73%	0.87%
Decision Table (search method: Greedy Stepwise	99.29%	4.10%	87.32%	0.56%
Decision Table (search method: Linear Forward Selection	99.65%	1.64%	87.32%	0.46%
Decision Table (search method: Rank Search	99.88%	0.00%	87.32%	0.77%
Decision Table (search method: ScatterSearchV1	98.47%	7.38%	87.01%	-0.46%
Decision Table (search method: Subset Size Forward Selection	98.70%	1.64%	86.49%	-0.98%
Multilayer Perceptron (1 hidden layer neurons = [number of attributes + number of classes]/2)	93.75%	36.07%	86.49%	-0.57%
Multilayer Perceptron (1 hidden layer neurons = number of attributes + number of classes)	93.87%	34.43%	86.39%	-0.26%
Multilayer Perceptron (1 hidden layer neurons = number of attributes)	95.87%	31.97%	87.84%	0.46%
Multilayer Perceptron (1 hidden layer 2 neurons)	93.99%	36.07%	86.70%	0.36%
PART (min number of instances/rule: 20)	98.23%	13.11%	87.53%	1.08%
PART (min number of instances/rule:10)	98.94%	8.20%	87.53%	0.98%
PART (min number of instances/rule:15)	99.06%	3.28%	87.01%	1.59%
PART (min number of instances/rule:2)	97.05%	18.85%	87.22%	0.56%
PART (min number of instances/rule:5)	96.70%	20.49%	87.11%	0.67%
Random Forest (10 Trees)	97.88%	14.75%	87.42%	-0.36%
Random Forest (2 Trees)	98.47%	11.48%	87.53%	1.69%
Random Forest (20 Trees)	96.46%	15.57%	86.29%	-1.70%
Random Forest (30 Trees)	98.58%	10.66%	87.53%	-0.36%
Random Forest (40 Trees)	98.70%	10.66%	87.63%	-0.05%
Random Forest (50 Trees)	98.58%	12.30%	87.73%	0.05%

Table 56: Results of several methods from second restricted by clinicians' version of Niguarda chronic dataset

METHOD	specificity	sensitivity	accuracy	IMPROVEMENT
Bayes Network	86.87%	46.38%	79.95%	1.98%
Decision Table	96.42%	7.25%	81.19%	-0.50%
Decision Table Naive Bayes Combination	91.34%	8.70%	77.23%	-4.70%
K Nearest Neighbors	82.09%	20.29%	71.53%	-4.95%
C 4.5	93.13%	17.39%	80.20%	4.70%
RIPPER	96.72%	17.39%	83.17%	4.21%
Multilayer Perceptron	89.55%	24.64%	78.47%	0.74%
Naive Bayes	82.09%	56.52%	77.72%	-2.72%
Non Nested Generalised Exemplars	93.43%	11.59%	79.46%	3.22%
PART	87.16%	27.54%	76.98%	-6.68%
Random Forest	95.22%	10.14%	80.69%	4.70%
Random Tree	84.78%	28.99%	75.25%	-5.45%
RBF Network	97.01%	4.35%	81.19%	0.00%
Voting Feature Intervals	75.82%	57.97%	72.77%	0.50%

Table 57: Results of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from second restricted by clinicians' version of Niguarda chronic dataset

METHOD	specificity	sensitivity	accuracy	IMPROVEMENT
Bayes Network (method for searching network structures: Global Tabu Search	91.64%	23.19%	79.95%	0.25%
Bayes Network (method for searching network structures: Global Hill Climber	91.94%	23.19%	80.20%	0.50%
Bayes Network (method for searching network structures: gK2	86.87%	46.38%	79.95%	1.98%
Bayes Network (method for searching network structures: Global Repeated Hill Climber	91.94%	23.19%	80.20%	0.50%
Bayes Network (method for searching network structures: ICS Search Algorithm	88.66%	37.68%	79.95%	1.73%
Bayes Network (method for searching network structures: Local Hill Climber	87.16%	33.33%	77.97%	1.98%
Bayes Network (method for searching network structures: lK2	86.87%	46.38%	79.95%	1.98%
Bayes Network (method for searching network structures: Local LAGD Hill Climber	87.16%	34.78%	78.22%	1.98%
Bayes Network (method for searching network structures: Local Repeated Hill Climber	87.16%	33.33%	77.97%	1.98%
Bayes Network (method for searching network structures: Local Simulated Annealing	93.43%	21.74%	81.19%	81.19%
Bayes Network (method for searching network structures: Local Tabu Search	86.27%	43.48%	78.96%	1.98%
Bayes Network (method for searching network structures: Local TAN	91.64%	24.64%	80.20%	-1.24%
Bayes Network (method for searching network structures: Naive Bayes	86.87%	46.38%	79.95%	1.98%
C 4.5 (min number of instances/leaf: 10)	100.00%	1.45%	83.17%	0.49%
C 4.5 (min number of instances/leaf: 15)	100.00%	0.00%	82.92%	0.50%
C 4.5 (min number of instances/leaf: 2)	97.61%	7.25%	82.18%	-0.99%
C 4.5 (min number of instances/leaf: 20)	100.00%	0.00%	82.92%	0.00%
C 4.5 (min number of instances/leaf: 5)	97.61%	4.35%	81.68%	-1.98%
Decision Table (search method: Best First	96.42%	7.25%	81.19%	-0.74%
Decision Table (search method: Greedy Stepwise	99.40%	2.90%	82.92%	0.00%
Decision Table (search method: Linear Forward Selection	98.21%	4.35%	82.18%	1.73%
Decision Table (search method: Rank Search	95.22%	15.94%	81.68%	-1.24%
Decision Table (search method: ScatterSearchV1	98.51%	2.90%	82.18%	-0.74%
Decision Table (search method: Subset Size Forward Selection	99.70%	2.90%	83.17%	0.00%
Multilayer Perceptron (1 hidden layer neurons = [number of attributes + number of classes]/2)	89.55%	24.64%	78.47%	-0.50%
Multilayer Perceptron (1 hidden layer neurons = number of attributes + number of classes)	88.06%	30.43%	78.22%	-2.72%
Multilayer Perceptron (1 hidden layer neurons = number of attributes)	87.46%	31.88%	77.97%	-2.23%
Multilayer Perceptron (1 hidden layer 2 neurons)	87.16%	24.64%	76.49%	-4.21%
PART (min number of instances/rule: 20)	97.61%	4.35%	81.68%	0.74%
PART (min number of instances/rule:10)	98.81%	1.45%	82.18%	0.00%
PART (min number of instances/rule:15)	92.24%	23.19%	80.45%	-0.50%
PART (min number of instances/rule:2)	98.21%	1.45%	81.68%	-0.74%
PART (min number of instances/rule:5)	96.12%	20.29%	83.17%	0.25%
Random Forest (10 Trees)	95.22%	10.14%	80.69%	-2.97%
Random Forest (2 Trees)	92.24%	15.94%	79.21%	0.50%
Random Forest (20 Trees)	97.61%	13.04%	83.17%	-1.24%

METHOD	specificity	sensitivity	accuracy	IMPROVEMENT
Random Forest (30 Trees)	96.72%	11.59%	82.18%	-1.24%
Random Forest (40 Trees)	97.91%	13.04%	83.42%	-1.24%
Random Forest (50 Trees)	98.21%	11.59%	83.42%	-0.25%

Table 58: Results of several methods from second restricted by clinicians' version of AMI dataset using SMOTE

METHOD	specificity	sensitivity	accuracy	IMPROVEMENT
Bayes Network	90.57%	91.15%	90.86%	-1.03%
C 4.5	95.64%	88.08%	91.86%	-0.02%
Decision Table	98.00%	86.42%	92.21%	-0.67%
Decision Table Naive Bayes Combination	92.81%	90.20%	91.50%	-0.50%
K Nearest Neighbors	93.99%	70.96%	82.48%	-9.35%
Multilayer Perceptron	91.75%	90.91%	91.33%	-0.44%
Naive Bayes	89.98%	90.44%	90.21%	-1.62%
Non Nested Generalised Exemplars	96.82%	80.64%	88.73%	-2.39%
PART	91.86%	89.02%	90.44%	-2.32%
Random Forest	96.82%	88.90%	92.86%	-0.08%
Random Tree	89.27%	90.32%	89.79%	0.26%
RBF Network	93.63%	87.96%	90.80%	-1.79%
RIPPER	98.47%	86.42%	92.45%	-0.67%
Voting Feature Intervals	98.23%	87.13%	92.68%	0.04%

Table 59: McNemar of several methods from second restricted by clinicians' version of AMI dataset using SMOTE

	Decision Table Naive Bayes Combination	Decision Table	C 4.5	RIPPER	Multilayer Perceptron	Naive Bayes	PART	RBF Network	Random Forest	Voting Feature Intervals	Bayes Network
Decision Table Naive Bayes Combination	NS	S	S	NS	S	NS	S	NS	S	S	NS
Decision Table	S	NS	S	S	S	S	S	S	S	NS	S
C 4.5	S	S	NS	S	S	S	S	S	S	S	S
RIPPER	NS	S	S	NS	S	S	S	S	S	NS	NS
Multilayer Perceptron	S	S	S	S	NS	S	S	S	NS	S	S
Naive Bayes	NS	S	S	S	S	NS	S	NS	S	S	S
PART	S	S	S	S	S	S	NS	S	S	S	S
RBF Network	NS	S	S	S	S	NS	S	NS	S	S	NS
Random Forest	S	S	S	S	NS	S	S	S	NS	S	S
Voting Feature Intervals	S	NS	S	NS	S	S	S	S	S	NS	S
Bayes Network	NS	S	S	NS	S	S	S	NS	S	S	NS

Table 60: Results of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from second restricted by clinicians' version of Niguarda AMI dataset using SMOTE

METHOD	specificity	sensitivity	accuracy	IMPROVEMENT
Bayes Network (method for searching network structures: Global Tabu Search	91.86%	90.32%	91.09%	91.09%
Bayes Network (method for searching network structures: Global Hill Climber	92.22%	90.55%	91.39%	-1.26%
Bayes Network (method for searching network structures: gK2	90.57%	91.15%	90.86%	-1.03%
Bayes Network (method for searching network structures: Global Repeated Hill Climber	92.22%	90.55%	91.39%	91.39%
Bayes Network (method for searching network structures: Local Hill Climber	90.45%	91.15%	90.80%	-0.91%
Bayes Network (method for searching network structures: lK2	90.57%	91.15%	90.86%	-1.03%
Bayes Network (method for searching network structures: Local LAGD Hill Climber	90.21%	91.15%	90.68%	-1.09%
Bayes Network (method for searching network structures: Local Repeated Hill Climber	90.45%	91.15%	90.80%	-0.91%
Bayes Network (method for searching network structures: Local Tabu Search	90.45%	91.15%	90.80%	-0.91%
Bayes Network (method for searching network structures: Local TAN	95.87%	88.43%	92.15%	-0.20%
Bayes Network (method for searching network structures: Naive Bayes	90.57%	91.15%	90.86%	-1.03%
Decision Table (search method: Best First	98.00%	86.42%	92.21%	0.15%
Decision Table (search method: Greedy Stepwise	98.11%	86.30%	92.21%	0.09%
Decision Table (search method: Linear Forward Selection	98.00%	86.07%	92.04%	-0.38%
Decision Table (search method: Rank Search	96.11%	87.72%	91.92%	-0.20%
Decision Table (search method: ScatterSearchV1	97.88%	87.01%	92.45%	-0.14%
Decision Table (search method: Subset Size Forward Selection	98.11%	86.66%	92.39%	-0.49%
C 4.5 (min number of instances/leaf: 10)	95.05%	88.55%	91.80%	-0.97%
C 4.5 (min number of instances/leaf: 15)	93.63%	89.14%	91.39%	-1.50%
C 4.5 (min number of instances/leaf: 2)	96.34%	88.43%	92.39%	0.15%
C 4.5 (min number of instances/leaf: 20)	93.51%	89.02%	91.27%	-1.03%
C 4.5 (min number of instances/leaf: 5)	95.64%	87.84%	91.74%	-1.32%
Multilayer Perceptron (1 hidden layer neurons = [number of attributes + number of classes]/2)	91.75%	90.91%	91.33%	-0.44%
Multilayer Perceptron (1 hidden layer neurons = number of attributes + number of classes)	91.98%	90.67%	91.33%	-0.97%
Multilayer Perceptron (1 hidden layer neurons = number of attributes)	93.40%	90.79%	92.09%	0.33%
Multilayer Perceptron (1 hidden layer 2 neurons)	92.69%	90.91%	91.80%	-0.97%
PART (min number of instances/rule:10)	95.52%	87.72%	91.62%	-0.61%
PART (min number of instances/rule:15)	94.22%	88.19%	91.21%	-0.79%
PART (min number of instances/rule:2)	94.58%	88.43%	91.50%	-0.55%
PART (min number of instances/rule: 20)	93.99%	88.78%	91.39%	-0.79%
PART (min number of instances/rule:5)	95.52%	88.90%	92.21%	-0.08%
Random Forest (10 Trees)	96.82%	88.90%	92.86%	-0.08%
Random Forest (2 Trees)	95.05%	88.55%	91.80%	-0.08%
Random Forest (20 Trees)	97.05%	88.55%	92.80%	-0.37%
Random Forest (30 Trees)	97.05%	88.43%	92.74%	-0.43%
Random Forest (40 Trees)	97.05%	88.67%	92.86%	-0.26%
Random Forest (50 Trees)	97.05%	88.55%	92.80%	-0.43%

Table 61: McNemar of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from second restricted by clinicians' version of Niguarda AMI dataset using SMOTE

	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest	Bayes Network
Decision Table	NS	NS	S	NS	S	NS
C 4.5	NS	NS	S	S	S	NS
Multilayer Perceptron	S	S	NS	S	NS	S
PART	NS	S	S	NS	S	NS
Random Forest	S	S	NS	S	NS	S
Bayes Network	NS	NS	S	NS	S	NS

Table 62: Results of several methods from second restricted by clinicians' version of chronic dataset using SMOTE

METHOD	specificity	sensitivity	accuracy	IMPROVEMENT
Bayes Network	87.66%	87.30%	87.48%	-1.78%
C 4.5	91.46%	86.03%	88.75%	0.18%
Decision Table	91.14%	76.19%	83.68%	-4.34%
Decision Table Naive Bayes Combination	89.87%	83.49%	86.69%	-2.02%
K Nearest Neighbors	79.43%	77.14%	78.29%	-3.53%
Multilayer Perceptron	87.66%	88.25%	87.96%	0.08%
Naive Bayes	83.86%	91.43%	87.64%	0.04%
Non Nested Generalised Exemplars	96.84%	53.02%	74.96%	-3.14%
PART	87.66%	85.71%	86.69%	-0.92%
Random Forest	92.72%	84.44%	88.59%	-1.36%
Random Tree	87.03%	85.08%	86.05%	-1.14%
RBF Network	88.61%	81.59%	85.10%	-3.46%
RIPPER	93.35%	83.81%	88.59%	-0.67%
Voting Feature Intervals	98.10%	74.92%	86.53%	-0.66%

Table 63: McNemar of several methods from second restricted by clinicians' version of chronic dataset using SMOTE

	Decision Table Naive Bayes Combination	Decision Table	C 4.5	RIPPER	Multilayer Perceptron	Naive Bayes	PART	Random Forest	Random Tree	Voting Feature Intervals	Bayes Network
Decision Table Naive Bayes Combination	NS	NS	NS	NS	S	S	S	S	S	NS	NS
Decision Table	NS	NS	NS	NS	S	S	S	S	S	NS	S
C 4.5	NS	NS	NS	S	S	S	S	S	S	S	S
RIPPER	NS	NS	S	NS	S	NS	S	S	S	NS	NS
Multilayer Perceptron	S	S	S	S	NS	S	NS	S	NS	S	S
Naive Bayes	S	S	S	NS	S	NS	S	S	S	NS	NS
PART	S	S	S	S	NS	S	NS	S	S	S	S
Random Forest	S	S	S	S	S	S	S	NS	NS	S	S
Random Tree	S	S	S	S	NS	S	S	NS	NS	S	S
Voting Feature Intervals	NS	NS	S	NS	S	NS	S	S	S	NS	NS
Bayes Network	NS	S	S	NS	S	NS	S	S	S	NS	NS

Table 64: Results of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from second restricted by clinicians' version of Niguarda chronic dataset using SMOTE

METHOD	specificity	sensitivity	accuracy	IMPROVEMENT
Bayes Network (method for searching network structures: Global Tabu Search	89.24%	86.98%	88.11%	-1.69%
Bayes Network (method for searching network structures: Global Hill Climber	89.24%	86.98%	88.11%	-2.11%
Bayes Network (method for searching network structures: gK2	87.66%	87.30%	87.48%	-1.78%
Bayes Network (method for searching network structures: Global Repeated Hill Climber	89.24%	86.98%	88.11%	-2.11%
Bayes Network (method for searching network structures: Local Hill Climber	86.39%	86.67%	86.53%	-2.73%
Bayes Network (method for searching network structures: lK2	87.66%	87.30%	87.48%	-1.78%
Bayes Network (method for searching network structures: Local LAGD Hill Climber	86.39%	86.67%	86.53%	-2.73%
Bayes Network (method for searching network structures: Local Repeated Hill Climber	86.39%	86.67%	86.53%	-2.73%
Bayes Network (method for searching network structures: Local Tabu Search	86.39%	86.67%	86.53%	-2.73%
Bayes Network (method for searching network structures: Local TAN	90.19%	88.25%	89.22%	-1.00%
Bayes Network (method for searching network structures: Naive Bayes	87.66%	87.30%	87.48%	-1.78%
Decision Table (search method: Best First	91.14%	76.19%	83.68%	-4.62%
Decision Table (search method: Greedy Stepwise	93.35%	75.24%	84.31%	-3.84%
Decision Table (search method: Linear Forward Selection	93.35%	77.14%	85.26%	-3.72%
Decision Table (search method: Rank Search	89.87%	84.44%	87.16%	-0.99%
Decision Table (search method: ScatterSearchV1	92.72%	78.10%	85.42%	-3.15%
Decision Table (search method: Subset Size Forward Selection	92.09%	75.24%	83.68%	-4.34%
C 4.5 (min number of instances/leaf: 10)	93.04%	83.49%	88.27%	0.26%
C 4.5 (min number of instances/leaf: 15)	89.56%	84.13%	86.85%	-1.72%
C 4.5 (min number of instances/leaf: 2)	93.35%	83.49%	88.43%	-0.69%
C 4.5 (min number of instances/leaf: 20)	85.44%	85.71%	85.58%	-4.09%
C 4.5 (min number of instances/leaf: 5)	89.87%	83.49%	86.69%	-2.71%
Multilayer Perceptron (1 hidden layer neurons = [number of attributes + number of classes]/2)	87.66%	88.25%	87.96%	0.08%
Multilayer Perceptron (1 hidden layer neurons = number of attributes)	88.92%	86.35%	87.64%	-0.52%
Multilayer Perceptron (1 hidden layer 2 neurons)	84.18%	87.62%	85.90%	-1.71%
Multilayer Perceptron (1 hidden layer neurons = number of attributes + number of classes)	88.92%	86.98%	87.96%	-0.47%
PART (min number of instances/rule:10)	89.87%	86.35%	88.11%	-0.73%
PART (min number of instances/rule:15)	88.61%	85.71%	87.16%	-0.72%
PART (min number of instances/rule:2)	92.72%	85.40%	89.07%	0.50%
PART (min number of instances/rule: 20)	86.39%	85.08%	85.74%	-2.28%
PART (min number of instances/rule:5)	90.82%	85.08%	87.96%	-1.16%
Random Forest (10 Trees)	92.72%	84.44%	88.59%	-1.36%
Random Forest (2 Trees)	93.04%	83.49%	88.27%	0.94%
Random Forest (20 Trees)	93.35%	85.08%	89.22%	-1.13%
Random Forest (30 Trees)	93.35%	84.44%	88.91%	-1.73%
Random Forest (40 Trees)	93.35%	83.81%	88.59%	-1.91%
Random Forest (50 Trees)	93.04%	83.81%	88.43%	-2.06%

Table 65: McNemar of random forest, c 4.5 part, multilayer perceptron and bayes network using different parameter values from second restricted by clinicians' version of Niguarda chronic dataset using SMOTE

	Decision Table	C 4.5	Multilayer Perceptron	PART	Random Forest	Bayes Network
Decision Table	NS	NS	S	NS	S	NS
C 4.5	NS	NS	S	NS	S	NS
Multilayer Perceptron	S	S	NS	S	S	S
PART	NS	NS	S	NS	S	NS
Random Forest	S	S	S	S	NS	S
Bayes Network	NS	NS	S	NS	S	NS

As expected the results when the datasets are balanced using SMOTE have higher sensitivity. The rules of the PART algorithm that were presented to the clinicians were satisfying and their clinical interpretation is analysed in the next chapter.

As future work in the Niguarda dataset the stratified balanced datasets methods must be tested in order to see if the accuracy and the rules are more satisfying. Moreover, the clinicians will have to check the rest of rule based classifiers in order to assure that the rules produced by PART algorithm are the ones that will be followed. Finally, the missing values must be treated in both Niguarda and Gissi dataset.

5.5 Clinicians feedback

In this chapter the interpretation of the most important classifiers is provided. These classifiers were proposed by the clinicians and after most of the initial results presented above were proved to be inaccurate and useless as part of a decision support system. So actually clinicians provided the feature subset selection that enabled data mining engineers to build classifiers that could actually extract new knowledge and be useful in a decision support system.

The clinicians have characterized the rules using the following categories:

- Red: rules at odds with common knowledge
- Green: rules that are in agreement with common knowledge and do not add new insights
- Grey: potentially new and interesting findings

Only the rules that were more accurate than the actual class distribution (i.e. above 91.75% for patients not developing late onset HF, i.e. class 0 and above 8.75% for patients that did develop late onset HF, i.e. class 1) were taken into consideration. This criterion was used as the main metric because it improves the accuracy of the prediction when compared to a random prediction which is represented from the class distribution in the real data set.

Retrospectively enrolled real world heart failure patients (with various types of AMI) – GISSI Data

The following table presents the results of the classifier that is based on the variables Diabetes, Ejection Fraction, AMI (acute myocardial infarction).These are considered as the three more important variables that evidently affect the presence of late onset heart failure.

Diabetes Ejection Fraction AMI

Samples	1224
Patients that did not develop late onset heart failure	1123 (91.75%)
Patients that developed late onset heart failure	101 (8.25 %)

Rule	Class	Samples following the rule	Correct	Wrong	Rule Accuracy
Diabetes = 0 AND EjectionFraction > 43.38 AND AMI = 2 AND EjectionFraction < 52	1	41	9	32	21,95%
Diabetes = 1	0	375	350	25	93,33%
AMI = 2	0	503	467	36	92,84%
AMI = 3	0	64	58	6	90,63%
AMI = 4 AND EjectionFraction > 43.898536 AND EjectionFraction < 68.599761 AND EjectionFraction < 60.0048	0	48	42	6	87,50%
AMI = 1 AND EjectionFraction > 58	1	74	10	64	13,51%
AMI = 4 AND EjectionFraction > 43.898536 AND EjectionFraction < 68	1	68	10	58	14,71%
AMI = 1 AND EjectionFraction > 29.45 AND EjectionFraction < 37	1	40	2	38	5,00%
AMI = 1 AND EjectionFraction > 50	1	147	17	130	11,56%
AMI = 1 AND EjectionFraction > 24.78 AND EjectionFraction < 48	1	143	15	128	10,49%

This classifier actually shows that none of the three commonly accepted as key indicators by the clinicians can provide a decision by itself. Especially Diabetes and AMI seem to provide the opposite results from what the clinicians have expected and at least in GISSI data set we can reach the conclusion that neither of them is a safe indicator in itself. AMI Site classification in GISSI data study is

1=inferoposterior;

2=anterior;

3=multiple;

4= not characterized by abnormal Q waves;

9=not definable

which means that 467 out of 503 subjects that had suffered anterior acute myocardial infarction were not readmitted to hospital, i.e. did not develop late on set heart failure, while only 36 were readmitted according to the following rule.

AMI = 2	0	503	467	36	92,84%

Of course many other factors may have contributed to this result but it is an undisputed fact extracted from our study and the specific classifier. Something similar happened with multiple AMI with 58 out of 64 subjects not being readmitted to the hospital.

AMI = 3	0	64	58	6	90,63%

Another interesting result coming in contrast with common knowledge is the fact the 350 out of the 375 subjects that had Diabetes were not readmitted to the hospital.

Diabetes = 1	0	375	350	25	93,33%

Concerning the VPH2 Decision Support System this classifier can only be used if it is customized by the user (add/ remove rules functionality) since the decision support from the original set of rules seems to be inadequate.

The classifier that proved to be the most interesting and intriguing for VPH2 clinicians is the one that the indicated feature subset consisted of 2) Diabetes, Ejection Fraction, AMI, Biochemical; the aim was to assess what do lab data (i.e. Cholesterol (total, HDL), White Blood Cells, Fibrinogen, Creatinine, Uric acid, Glycaemia, PCR, SGOT / SGPT, Na, Triglycerides and Aematocrit) in general (and which one in particular) add in the predictive accuracy of late on set heart failure/ readmission to the hospital.

Diabetes Ejection Fraction AMI Biochemical

No Samples	1224
Healthy	1123 (91.75%)
Not Healthy	101 (8.25 %)

Rule	Class	Samples following the rule	Correct	Wrong	Rule Accuracy
Diabetes = 1 AND TotChol < 237.715419 AND Trigl < 129	0	122	117	5	95,90%
Diabetes = 1 AND TotChol < 237.715419 AND AMI = 2	0	107	97	10	90,65%
Diabetes = 1 AND TotChol > 237	0	59	59	0	100,00%
AMI = 4 AND SGOT > 15	0	144	137	7	95,14%
EjectionFraction > 40.98 AND Diabetes = 0 AND Trigl < 76 AND AMI = 1 AND UricAcid < 6	0	7	7	0	100,00%
EjectionFraction > 40.98 AND Diabetes = 0 AND Creatinine < 1.199417 AND Creatinine < 1.1 AND Creatinine <	1	0	0	0	--
EjectionFraction > 40.98 AND Diabetes = 0 AND AMI = 1 AND Glycaemia < 142 AND NA > 139.000696 AND Gly	1	22	5	17	22,73%
EjectionFraction < 40.98 AND PCR < 0.504345 AND Glycaemia < 111	0	81	81	0	100,00%
Trigl > 179.854835 AND Glycaemia < 89	0	84	83	1	98,81%
Trigl > 73 AND Diabetes = 1 AND SGPT < 73 AND NA > 140	0	131	122	9	93,13%
Trigl > 73 AND Glycaemia > 146 AND Diabetes = 0 AND Aematocrit < 42	1	1	1	0	100,00%
Trigl > 172.954682 AND TotChol < 261 AND hdlChol < 35.004969 AND NA < 144.394624 AND Aematocrit < 40	0	56	54	2	96,43%
Trigl > 73 AND Creatinine < 1.199417 AND Creatinine < 1.1 AND Creatinine < 0.998969 AND Creatinine > 0	1	340	31	309	9,12%
Trigl > 75 AND EjectionFraction < 31	0	82	79	3	96,34%
Trigl < 75	0	79	75	4	94,94%
AMI = 4 AND Aematocrit > 40	0	64	63	1	98,44%
NA < 134 AND Fibrinogen > 333	0	34	34	0	100,00%
Trigl > 239.275088 AND hdlChol < 40	0	87	83	4	95,40%
NA > 142.998533 AND Glycaemia < 80	0	27	27	0	100,00%
AMI = 1 AND Creatinine < 0.899822 AND Diabetes = 0 AND Creatinine > 0	1	31	8	23	25,81%
AMI = 3 AND Fibrinogen < 323	1	13	3	10	23,08%
AMI = 3 AND WhiteBloodcellcounts < 9	0	46	44	2	95,65%
AMI = 1 AND NA > 142.998533 AND Aematocrit < 45	0	93	89	4	95,70%
AMI = 1 AND PCR < 35.074973 AND NA > 139.000696 AND TotChol > 212.071814 AND Diabetes = 0 AND NA <	1	17	6	11	35,29%
Creatinine < 0.799829 AND Diabetes = 0 AND Creatinine > 0	1	37	6	31	16,22%
AMI = 4 AND PCR < 3.48237 AND SGPT < 19	0	23	22	1	95,65%
AMI = 2 AND Fibrinogen > 377.427873 AND Glycaemia > 82 AND NA > 137.483522 AND UricAcid < 6.841139 A	0	29	28	1	96,55%
AMI = 1 AND NA < 142 AND PCR < 35.074973 AND Creatinine < 1.000503 AND Trigl < 170 AND Creatinine > 0.	0	64	59	5	92,19%
AMI = 1 AND PCR > 30	1	15	5	10	33,33%
AMI = 1 AND NA > 142	1	106	5	101	4,72%
AMI = 1 AND Aematocrit < 34.992538 AND Fibrinogen > 307	1	41	8	33	19,51%
AMI = 1 AND Diabetes = 1 AND Creatinine > 0.817957 AND EjectionFraction < 48	1	28	3	25	10,71%
AMI = 1 AND SGPT > 63	0	51	46	5	90,20%
AMI = 1 AND Diabetes = 1 AND Trigl > 200	0	51	49	2	96,08%
AMI = 1 AND Diabetes = 1 AND WhiteBloodcellcounts < 6	1	21	3	18	14,29%
AMI = 1 AND Diabetes = 0	1	320	35	285	10,94%
AMI = 2 AND Fibrinogen > 221.371439 AND Fibrinogen < 269.596462 AND SGPT > 12	1	26	6	20	23,08%
AMI = 2 AND Fibrinogen > 270.493281 AND Glycaemia > 101 AND Creatinine > 1	1	80	9	71	11,25%
AMI = 2 AND TotChol > 232.923849 AND SGOT < 23	0	62	60	2	96,77%
AMI = 2 AND Creatinine > 1	0	261	246	15	94,25%
AMI = 2 AND Fibrinogen > 211.862427 AND UricAcid < 4.302888 AND PCR > 1	0	20	20	0	100,00%
AMI = 2 AND Fibrinogen > 211.862427 AND TotChol < 186.169345 AND SGOT > 21.288612 AND hdlChol < 47	1	51	8	43	15,69%
AMI = 2 AND PCR < 31 AND Fibrinogen > 274.216407 AND WhiteBloodcellcounts > 7.012166 AND SGPT > 36	0	79	75	4	94,94%
AMI = 2 AND Fibrinogen > 211.862427 AND PCR < 16	0	303	280	23	92,41%
Fibrinogen < 211	0	59	55	4	93,22%
AMI = 2	1	503	36	467	7,16%
AMI = 1	0	478	433	45	90,59%
AMI = 4	1	171	14	157	8,19%
AMI = 9	0	8	8	0	100,00%

In the following some of the rules coming in contrast with common knowledge and all the rules characterized from the clinicians as potentially new interesting findings are analyzed in order to present the knowledge that can be extracted from this classifier and the way a researcher should think when using the VPH2 decision support.

[redacted]

This rule supports the conclusion reached also with the first classifier: diabetes by itself is not enough to prognose late onset heart failure even though it still remains an important risk factor. As long as the total cholesterol and the triglycerides are below certain thresholds the prediction for not developing late on set heart failure is very accurate: 95.90%.

[redacted]

Again the total cholesterol threshold seems to be more important than AMI and familiar Diabetes. This rule cannot be part of the decision support since it is less accurate than the actual class distribution and this is why it was originally ignored by the clinicians that reviewed this classifier after the suggestion of the data mining engineers of course.

[redacted]

This is actually one of the most controversial yet potentially very interesting findings in GISSI data set. There are 59 cases that even if diabetes is present and the total cholesterol is above 237 (even though the average is 262), haven't developed late onset heart failure. One such case is for example a 55 years old woman which has normal features (normal values in most variables available in her file) and which has certain interesting characteristics: PCR value is normal (in most subjects it isn't), BMI is 33 and she has been treated with Beta-Blockers and PUFA. The utility of this rule is the comparison with a similar new case to be assessed. A new patient may be better treated if the clinician is aware of a success treatment story of a patient with similar characteristics in the past.

[redacted]

This rule combines information: when EF is normal, the patient doesn't have diabetes and the Uric Acid value is normal, patients that had suffered inferoposterior AMI and have low triglycerides most probably will not be readmitted. The issue is of course that in only 7 cases this rule is confirmed and thus the absolute accuracy it has may be disputed. This rule needs to be applied in independent data sets to prove its worth. In any case it is reasonable and potentially very useful for patients with inferoposterior AMI.

[redacted]

In this rule some parts are controversial: patients with normal or mild depression of EF, probably normal NA, abnormal aematocrit and glycaemia relatively above normal threshold will probably develop late onset heart failure. Again the subset of patients following this rule should be re-examined in order to understand if there is anything interesting in this small population.

[redacted]

Even though this rule is at odds with common knowledge, the average values of these specific variables in this population (consisting of the 52 subjects for which we know the EF) are close to normal: the average PCR is 0.43, the average glycaemia is 104 and the average EF is the strange finding since it is 33% (below the threshold that is 40%). Moreover, only 5 women are part of this population which may be not a coincidental fact and may worth having a second look at it.

Trigl > 179.854835 AND Glycaemia < 89	0	84	83	1	98,81%

This rule is impressively accurate. The average triglycerides are 247 and the average glycaemia is 81.2. The main characteristics of this population is that most subjects are males (only 7 females), smokers (with an average of 23 cigarettes per day and only 10 subjects not smoking), with an average age 56 years old, a quite normal BMI average around 27, and have suffered of various types of AMI.

Trigl > 73 AND Diabetes = 1 AND SGPT < 73 AND NA > 140	0	131	122	9	93,13%

This rule is rather controversial. Obviously it confirms that the normal triglycerides and NA values are very important factors for avoiding readmission to the hospital. In fact the average triglycerides is 160 and NA is 139,81 which both are within the normal ranges and prove themselves as more important factor compared to the presence of Diabetes. Of course this rule (and by the way all rules that define a lower threshold and not an upper) must be completed by the user in order to define normality values of the questioned lab exams.

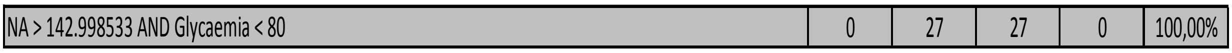
Trigl < 75	0	79	75	4	94,94%

The most interesting aspect of this rule, beyond its accuracy based in one parameter only, is that the population in which it is applied and confirmed is mixed. Some of the patients are diabetic; some of them have familiar hypertension; they take different drugs; they are about 62 years old; the average BMI is 26,6; most of them are smokers. SO this subset is worth of a more careful look in order to understand what other characteristics apart from low triglycerides that anyway indicate a healthy diet, may be preventive against late onset heart failure.

AMI = 4 AND Aematocrit > 40	0	64	63	1	98,44%

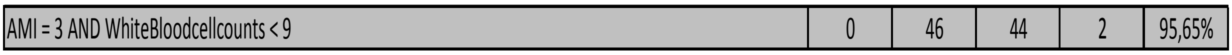
AMI = 4 AND PCR < 3.48237 AND SGPT < 19	0	23	22	1	95,65%

AMI = 4 is translated as AMI not characterized by abnormal Q waves and the average aematocrit in this subset is 44,28. These two rules suggest (in an impressively accurate manner) that people with that have suffered from this particular AMI will avoid readmission if they have a normal aematocrit and relatively low ASL results. With the exception of Total Cholesterol which is above normal range all the other characteristics of this group are normal.



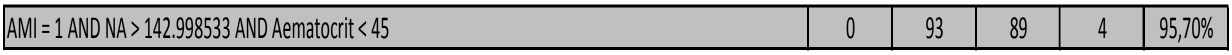
NA > 142.998533 AND Glycaemia < 80	0	27	27	0	100,00%

Again this rule is based on just two lab exams results: NA and blood glucose. The average NA is 145,55 which is slightly above normal NA which is 143. The average blood glucose is 73, which in fact is close to lower limit (70). All other average values of the lab exams in this relatively restricted population (27 samples) are normal with the exception of PCR. Another interesting characteristic is the BMI average which is 25,05, i.e. very close to the normal upper limit.



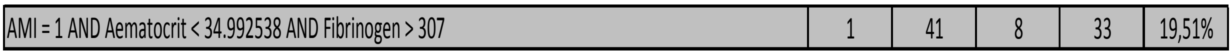
AMI = 3 AND WhiteBloodcellcounts < 9	0	46	44	2	95,65%

This rule is applied to a certain subset of patients that had suffered multiple AMI. For those patients the prognosis is optimistic as long as their white blood cell count examination is below 9.000. In fact the average in this certain population is 6.880. What is controversial in this population is the high fibrinogen and blood glucose values among these patients and the low aematocrit they present.



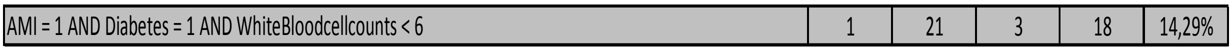
AMI = 1 AND NA > 142.998533 AND Aematocrit < 45	0	93	89	4	95,70%

As it happens in the previous rule too again this rule is applied to specific patients: those that were diagnosed with an inferoposterior AMI. For those patients the prognosis is optimistic as long as their lab exams and especially NA and aematocrit are normal.



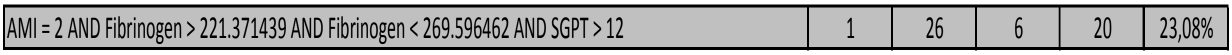
AMI = 1 AND Aematocrit < 34.992538 AND Fibrinogen > 307	1	41	8	33	19,51%

This rule makes a negative prognosis for the patients: It suggests that a low aematocrit for patients suffering from inferoposterior AMI means a higher possibility of readmission to the hospital. This population is rather older than previous (67,4 while most are around 62,5) and it presentes elevated fibrinogen and blood glucose levels, always in an average level, and a very low aematocrit average value at 31,7. These are obviosly the main factors contributing to a pesimistic prognosis for these patients. By the way the accuracy 19,51% is much better of the 8,25% that the class distribution between the dataset samples, presents. And this also stands for the following rules that predict class 1, i.e. patients that will develop late onset heart failure.

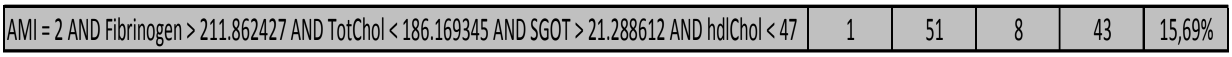


AMI = 1 AND Diabetes = 1 AND WhiteBloodcellcounts < 6	1	21	3	18	14,29%

This rule is complementary to the above one. It suggests that apart other factors the presence of diabetes and the white blood cells count can be predictive variables for the development of late onset HF. Of course the relatively few cases in which the rule is applied strengthen the conclusion that diabetes itself cannot be considered as a very strong risk factor.

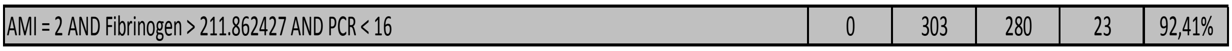


AMI = 2 AND Fibrinogen > 221.371439 AND Fibrinogen < 269.596462 AND SGPT > 12	1	26	6	20	23,08%



AMI = 2 AND Fibrinogen > 211.862427 AND TotChol < 186.169345 AND SGOT > 21.288612 AND hdlChol < 47	1	51	8	43	15,69%

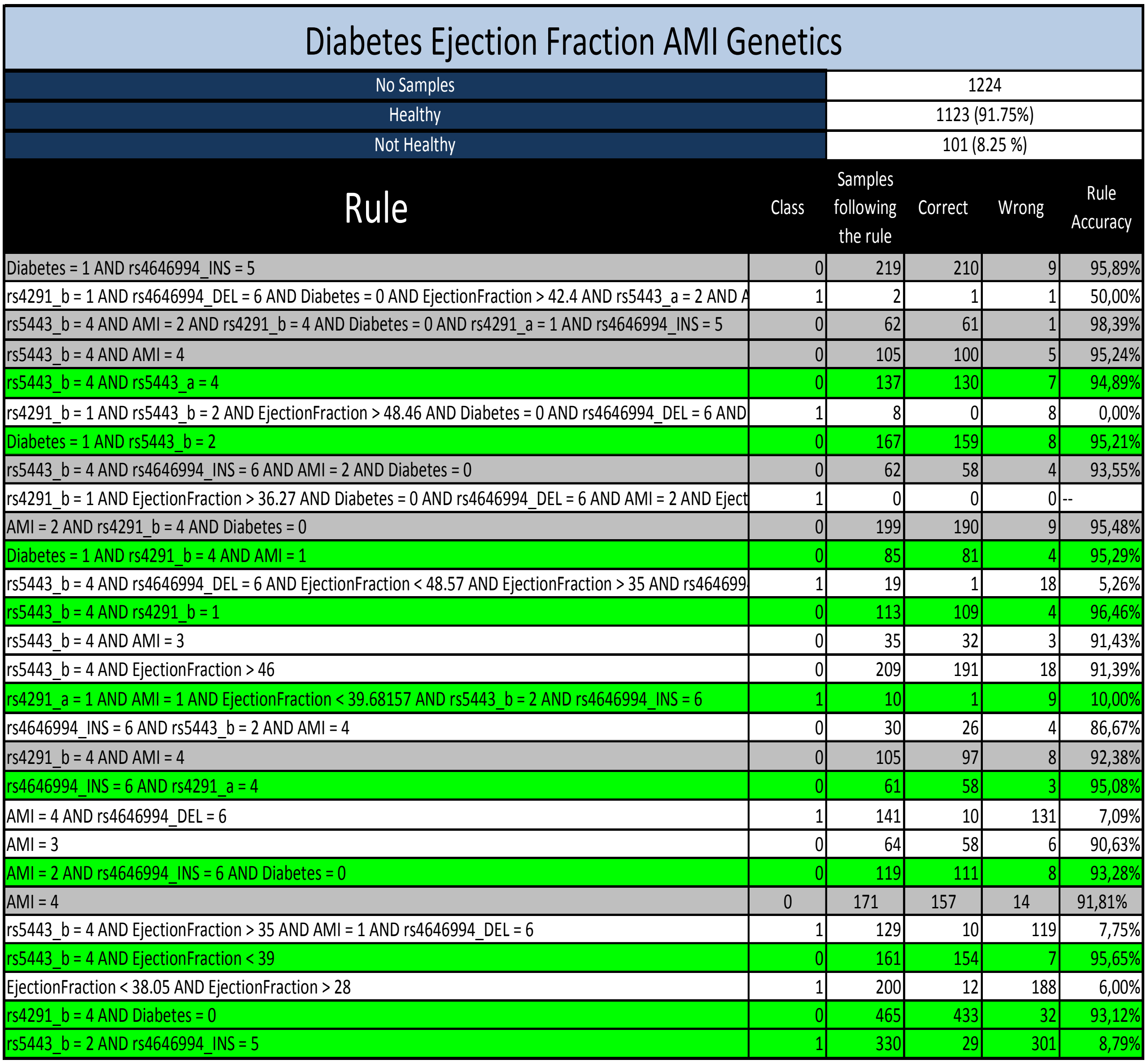
When combined these two rules imply that patients suffering from anterior AMI (504 in GISSI study) will be readmitted to the hospital even if the lab exams are normal or close to normality. What is notable is that most of these patients (above 70%) were treated with ACE inhibitors. Of course any rules supporting decisions for class 1 must be further investigated since the small amount of such samples and the resulting unbalanced dataset may be misleading when trying to reach any conclusions. However, they are indicative and potentially intriguing, for the clinical researchers results.



AMI = 2 AND Fibrinogen > 211.862427 AND PCR < 16	0	303	280	23	92,41%

This rule supports decisions for the patients suffering from anterior AMI. As noted above the rules for this subset is rather controversial and the conclusion is that the lab exams cannot provide adequate decision support. This is also due to the fact that most patients fall into this category. The only suggestion is that a low fibrinogen is associated with better prognosis for those patients.

Biologists feedback



Diabetes Ejection Fraction AMI Genetics					
No Samples		1224			
Healthy		1123 (91.75%)			
Not Healthy		101 (8.25 %)			
Rule	Class	Samples following the rule	Correct	Wrong	Rule Accuracy
Diabetes = 1 AND rs4646994_INS = 5	0	219	210	9	95,89%
rs4291_b = 1 AND rs4646994_DEL = 6 AND Diabetes = 0 AND EjectionFraction > 42.4 AND rs5443_a = 2 AND A	1	2	1	1	50,00%
rs5443_b = 4 AND AMI = 2 AND rs4291_b = 4 AND Diabetes = 0 AND rs4291_a = 1 AND rs4646994_INS = 5	0	62	61	1	98,39%
rs5443_b = 4 AND AMI = 4	0	105	100	5	95,24%
rs5443_b = 4 AND rs5443_a = 4	0	137	130	7	94,89%
rs4291_b = 1 AND rs5443_b = 2 AND EjectionFraction > 48.46 AND Diabetes = 0 AND rs4646994_DEL = 6 AND	1	8	0	8	0,00%
Diabetes = 1 AND rs5443_b = 2	0	167	159	8	95,21%
rs5443_b = 4 AND rs4646994_INS = 6 AND AMI = 2 AND Diabetes = 0	0	62	58	4	93,55%
rs4291_b = 1 AND EjectionFraction > 36.27 AND Diabetes = 0 AND rs4646994_DEL = 6 AND AMI = 2 AND Eject	1	0	0	0	--
AMI = 2 AND rs4291_b = 4 AND Diabetes = 0	0	199	190	9	95,48%
Diabetes = 1 AND rs4291_b = 4 AND AMI = 1	0	85	81	4	95,29%
rs5443_b = 4 AND rs4646994_DEL = 6 AND EjectionFraction < 48.57 AND EjectionFraction > 35 AND rs464699	1	19	1	18	5,26%
rs5443_b = 4 AND rs4291_b = 1	0	113	109	4	96,46%
rs5443_b = 4 AND AMI = 3	0	35	32	3	91,43%
rs5443_b = 4 AND EjectionFraction > 46	0	209	191	18	91,39%
rs4291_a = 1 AND AMI = 1 AND EjectionFraction < 39.68157 AND rs5443_b = 2 AND rs4646994_INS = 6	1	10	1	9	10,00%
rs4646994_INS = 6 AND rs5443_b = 2 AND AMI = 4	0	30	26	4	86,67%
rs4291_b = 4 AND AMI = 4	0	105	97	8	92,38%
rs4646994_INS = 6 AND rs4291_a = 4	0	61	58	3	95,08%
AMI = 4 AND rs4646994_DEL = 6	1	141	10	131	7,09%
AMI = 3	0	64	58	6	90,63%
AMI = 2 AND rs4646994_INS = 6 AND Diabetes = 0	0	119	111	8	93,28%
AMI = 4	0	171	157	14	91,81%
rs5443_b = 4 AND EjectionFraction > 35 AND AMI = 1 AND rs4646994_DEL = 6	1	129	10	119	7,75%
rs5443_b = 4 AND EjectionFraction < 39	0	161	154	7	95,65%
EjectionFraction < 38.05 AND EjectionFraction > 28	1	200	12	188	6,00%
rs4291_b = 4 AND Diabetes = 0	0	465	433	32	93,12%
rs5443_b = 2 AND rs4646994_INS = 5	1	330	29	301	8,79%

Here we discuss the most prominent data mining results were genetic parameters were used. As well we used the following classification: only the rules that were more accurate than the actual class distribution (i.e. above 91.75% for patients not developing late onset HF, i.e. class 0 and above 8.75% for patients that did develop late onset HF, i.e. class 1) were taken into consideration. The three genetic variants used (rs4291, rs5443 and rs4646994) were shown to be associated with late onset HF in D4.2 (all p-values <

0.05). More precisely, we identified a significant association for two genes within the study population. One gene encodes for the angiotensin I-converting enzyme (ACE), the other for the guanine nucleotide-binding protein (GNB3). Two genetic variations positioned in ACE, termed rs4291_a=1 and rs4646994_INS=6 and one positioned in GNB3, termed rs5443_b=2 marked the two identified genes. This translates as follows. The three alleles of the identified variants, namely rs4291=1, rs4646994=6 and rs5443=2 are predictors for late-onset HF in the study population used. The other alleles rs4291=4, rs4646994=5 and rs5443=4 are not associated with late-onset HF. Neither of the variants used are predictors for MI since they were not associated with MI. It has to be underlined that the functionality of the variants identified has not been experimentally proven. Since we cannot be sure how the present allele effects the function of the protein, all findings are marked in grey. Some findings may not be in agreement with common knowledge, since genetic data is combined with biochemical and other markers. Discrepancies might point towards underlying unknown mechanisms and present potential starting points for selective research activities

Based on the variables Diabetes, Ejection fraction, AMI and Genetics we found:

Diabetes = 1 AND rs4646994_INS = 5	0	219	210	9	95,89%

Despite the finding that Diabetes is not a predictor for late-onset HF in this population, rs4646994=5 marks patients who did not develop late-onset HF. Remarkably, Diabetes alone reaches an accuracy of 93.3%. Addition of total cholesterol and triglycerides raise the rule accuracy to 95.9%. The same effect is observed for the genetic information rs4646994=5. This genetic marker elevates accuracy of risk prediction by the same extend as the biochemical marker.

AMI = 4	0	171	157	14	91,81%

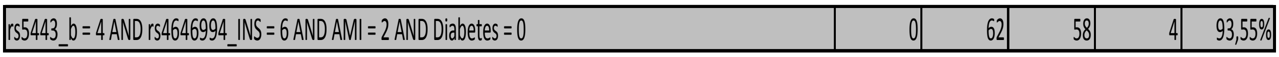
rs4291_b = 4 AND AMI = 4	0	105	97	8	92,38%

rs5443_b = 4 AND AMI = 4	0	105	100	5	95,24%

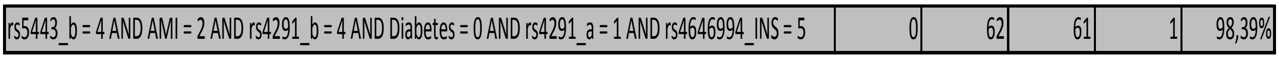
A quite interesting finding. 91.8% of patients with AMI status 4 were not readmitted to the hospital. Addition of the genetic information rs4291=4 raised the rule accuracy to 92.4% while rs5443=4 raised the accuracy to 95.2%, suggesting a higher predictive value of variant rs5443=4. In both cases exactly 105 samples followed this rule.

AMI = 2	0	503	467	36	92,84%

AMI = 2 AND rs4291_b = 4 AND Diabetes = 0	0	199	190	9	95,48%

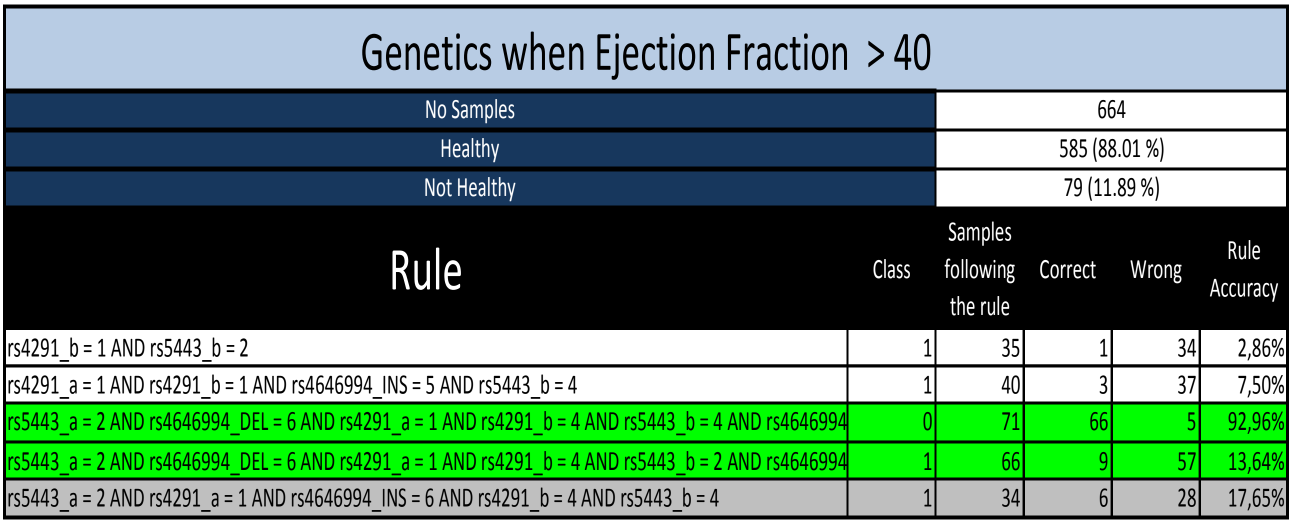


rs5443_b = 4 AND rs4646994_INS = 6 AND AMI = 2 AND Diabetes = 0	0	62	58	4	93,55%



rs5443_b = 4 AND AMI = 2 AND rs4291_b = 4 AND Diabetes = 0 AND rs4291_a = 1 AND rs4646994_INS = 5	0	62	61	1	98,39%

Looking at these four rules we identify again that patients who had suffered from anterior acute myocardial infarction were not readmitted to the hospital. The rule accuracy is raised to 95.5% by adding rs4291=4 and Diabetes=0, while diabetes was identified not to be a risk predictor for late-onset HF and vice versa. Combining this rule with rs4646994=6, which is associated with late-onset HF, lowers the rule accuracy to 93.5%. Combining all “protective” alleles of the three genetic variants in one rule, raises rule accuracy to 98.4%, which is a difference of 5.6%, even if rs4291 is heterozygous (rs4291=4 and rs4291=1). This is a good example that combination of genetic variants can remarkably increase accuracy of outcome prediction.

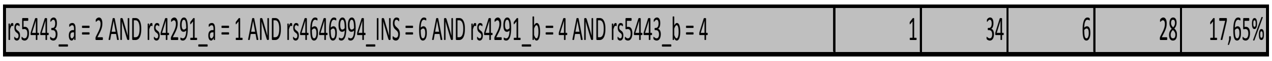


Genetics when Ejection Fraction > 40

No Samples	664
Healthy	585 (88.01 %)
Not Healthy	79 (11.89 %)

Rule	Class	Samples following the rule	Correct	Wrong	Rule Accuracy
rs4291_b = 1 AND rs5443_b = 2	1	35	1	34	2,86%
rs4291_a = 1 AND rs4291_b = 1 AND rs4646994_INS = 5 AND rs5443_b = 4	1	40	3	37	7,50%
rs5443_a = 2 AND rs4646994_DEL = 6 AND rs4291_a = 1 AND rs4291_b = 4 AND rs5443_b = 4 AND rs4646994	0	71	66	5	92,96%
rs5443_a = 2 AND rs4646994_DEL = 6 AND rs4291_a = 1 AND rs4291_b = 4 AND rs5443_b = 2 AND rs4646994	1	66	9	57	13,64%
rs5443_a = 2 AND rs4291_a = 1 AND rs4646994_INS = 6 AND rs4291_b = 4 AND rs5443_b = 4	1	34	6	28	17,65%

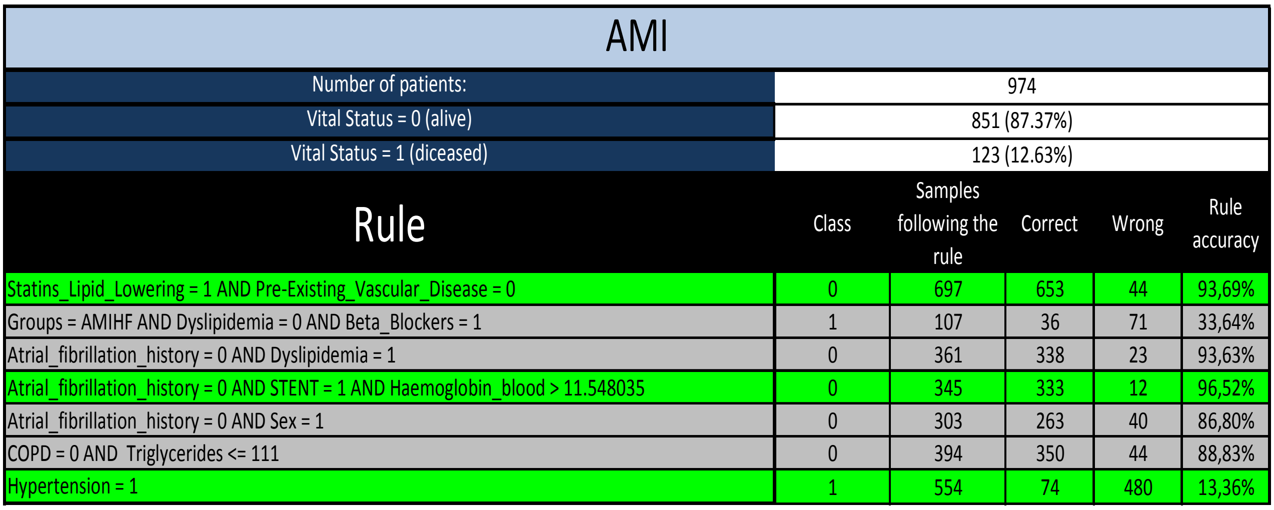
Notably, high rule accuracy was often observed in prediction of positive outcomes (no late-onset HF). Based on the variables “Genetics when Ejection fraction > 40” we identified the following rule:



rs5443_a = 2 AND rs4291_a = 1 AND rs4646994_INS = 6 AND rs4291_b = 4 AND rs5443_b = 4	1	34	6	28	17,65%

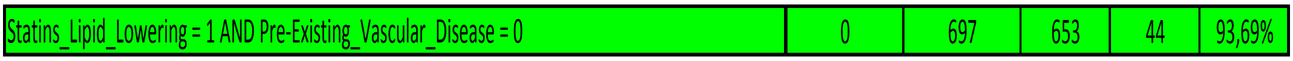
Patients with the combined genetic markers associated with late onset HF, rs4291=1, rs4646994=6 and rs5443=2, even if they are present in a heterozygous situation, are more likely to be readmitted to the hospital. This could mark rs4291=1 and rs5443=2 as risk alleles with higher impact on the possible outcome than the protective effect of the rs4291=4 and rs5443=4. The effect over average is 5.7% which exactly resembles the combined effect of the three alleles not associated with late-onset HF in the dataset based on the variables Diabetes, Ejection fraction, AMI and Genetics. **Retrospectively enrolled real world ischemic heart disease patients - NIGUARDA Data**

In the following the initial results of the application of data mining methods with the aim to derive rules that improve clinician decision-making based on NIGUARDA dataset and more specifically the AMI subset are presented. The clinical interpretation is also provided.



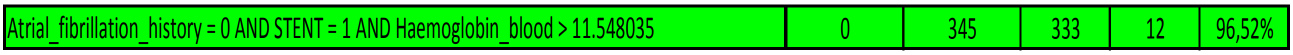
AMI					
Number of patients:	974				
Vital Status = 0 (alive)	851 (87.37%)				
Vital Status = 1 (diceased)	123 (12.63%)				
Rule	Class	Samples following the rule	Correct	Wrong	Rule accuracy
Statins_Lipid_Lowering = 1 AND Pre-Existing_Vascular_Disease = 0	0	697	653	44	93,69%
Groups = AMIHF AND Dyslipidemia = 0 AND Beta_Blockers = 1	1	107	36	71	33,64%
Atrial_fibrillation_history = 0 AND Dyslipidemia = 1	0	361	338	23	93,63%
Atrial_fibrillation_history = 0 AND STENT = 1 AND Haemoglobin_blood > 11.548035	0	345	333	12	96,52%
Atrial_fibrillation_history = 0 AND Sex = 1	0	303	263	40	86,80%
COPD = 0 AND Triglycerides <= 111	0	394	350	44	88,83%
Hypertension = 1	1	554	74	480	13,36%

Rules were classified as defined above. Results are consistent in general with indications from the literature even in the reperfusion and statin era. The following are two examples



Statins_Lipid_Lowering = 1 AND Pre-Existing_Vascular_Disease = 0	0	697	653	44	93,69%

The high accuracy of this rule in the prediction of a good outcome in this wide population subset confirms results from RCT on secondary prevention with statins in patients who do not have coexistent vascular disease in district other than the coronary one. The following rule also confirms results of previous studies [16;17;19;51]



Atrial_fibrillation_history = 0 AND STENT = 1 AND Haemoglobin_blood > 11.548035	0	345	333	12	96,52%

The negative prognostic impact of hypertension had been previously described in the classical Cox model from the GISSI Prevenzione dataset [10], and is confirmed by our results, even with a relatively low accuracy.

Although the predictive role of clinical HF on presentation and atrial fibrillation are well established in AMI, the combination with other predictors is novel and intriguing; in particular prescription of beta-blockers in this subset when still unstable is suggested by the negative impact of this class of drugs of proven efficacy in heart failure.

Data mining appears to provide additional prognostic insight when compared to Cox multivariable models

Retrospective chronic ischemic heart disease (cIHD) and chronic ischemic heart failure (cIHF) patients – NIGUARDA Data

In the following the initial results of the application of data mining methods with the aim to derive rules that improve clinician decision-making based on NIGUARDA dataset and more specifically the cIHD and cIHF subsets are presented. The clinical interpretation is also provided.

cIHD or cIHF					
Number of patients:	404				
Vital Status = 0 (alive)	335 (82.92%)				
Vital Status = 1 (diceased)	69 (17.03%)				
Rule	class	Samples following the rule	correct	wrong	accuracy
Statins_Lipid_Lowering = 1 AND Groups = cIHD	0	133	129	4	96,99%
Statins_Lipid_Lowering = 0 AND Aldosterone_Antag. = 1 AND CABG_index_admission = 0 AND Smoking_Habits = 0	1	31	14	17	45,16%
loop_diuretics_dose <= 86.638455 AND Beta_Blockers = 1	0	271	246	25	90,77%
Number_bypass = 0 AND Calcium_Channel_Blockers = 0 AND COPD = 0 AND Previous_STENT = 1	0	73	67	6	91,78%
Number_bypass = 0 AND Calcium_Channel_Blockers = 0 AND Diabetes = 1	1	74	22	52	29,73%

Rules were classified as defined above. Results are consistent in general with indications from the literature. The following are two examples

Statins_Lipid_Lowering = 1 AND Groups = cIHD	0	133	129	4	96,99%

Statin treatment and the absence of heart failure are associated to a good outcome with very high accuracy

loop_diuretics_dose <= 86.638455 AND Beta_Blockers = 1	0	271	246	25	90,77%

Lower doses of loop-diuretics and administration (and consequently tolerability) of beta-blockers are also known to be associated to a better prognosis.

Statins_Lipid_Lowering = 0 AND Aldosterone_Antag. = 1 AND CABG_index_admission = 0 AND Smoking_Habits = 0	1	31	14	17	45,16%

This rule is somewhat inconsistent with common knowledge, but it has on the other hand a very poor accuracy and is applicable to a limited number of subjects. Conversely the last 2 rules are potentially interesting; calcium channel blockers are not a recommended therapy for ischemic heart disease unless beta-blockers are not tolerated; the poor outcome of diabetics when not revascularized is also known, but the interaction with CCB as potential specific treatment is intriguing

Number_bypass = 0 AND Calcium_Channel_Blockers = 0 AND COPD = 0 AND Previous_STENT = 1	0	73	67	6	91,78%
Number_bypass = 0 AND Calcium_Channel_Blockers = 0 AND Diabetes = 1	1	74	22	52	29,73%

REFERENCES

[1] J. P. Hellermann, S. J. Jacobsen, B. J. Gersch, G. S. Rodeheffer, and V. L. Reeder, "Heart failure after myocardial infarction: a review," *Am J Med*, vol. 113, pp. 324-340, 2002.

[2] L. Bolognese, A. N. Neskovic, G. Parodi, G. Cerisano, P. G. Buonamici, P. G. Santoro, and D. Antoniucci, "Left ventricular remodeling after primary coronary angioplasty Patterns of left ventricular dilation and long-term prognostic implications," *Circulation*, vol. 106, pp. 2351-2357, 2002.

[3] E. F. Lewis, L. A. Moye, J. L. Rouleau, F. M. Sacks, J. M. O. Arnold, J. W. Warnica, G. C. Flaker, E. Braunwald, and M. A. Pfeffer , "Predictors of late development of heart failure in stable survivors of myocardial infarction," *J Am Coll Cardiol*, vol. 42, pp. 1446-1453, 2003.

[4] J. P. Hellermann, S. J. Jacobsen, M. M. Redfield, V. L. Reeder, S. A. Weston, and V. L. Roger, "Heart failure after myocardial infarction: clinical presentation and survival," *Eur J Heart Fail*, vol. 7, pp. 119-125, 2005.

[5] GISSI-Prevenzione Investigators (Gruppo Italiano per lo Studio della Sopravvivenza nell'Infarto miocardico), "Dietary supplementation with n-3 polyunsaturated fatty acids and vitamin E after myocardial infarction: results of the GISSI-Prevenzione trial," *Lancet*, vol. 354, no. 9177, pp. 447-455, 1999.

[6] J. B. De Kok, E. T. G. Wiegerinck, B. A. J. Giesendorf, and D. W. Swinkels, "Rapid Genotyping of Single Nucleotide Polymorphisms Using Novel Minor Groove Binding DNA Oligonucleotides (MGB Probes)," *Human Mutation*, vol. 19, no. 5, pp. 554-559, 2002.

[7] G. S. Bleumink, A. F. Schut, M. C. Sturkenboom, J. W. Deckers, C. M. van Duijn, and B. H. Stricker, "Genetic polymorphisms and heart failure," *Genet Med*, vol. 6, no. 6, pp. 465-474, 2004.

[8] "gPLINK," http://pngu.mgh.harvard.edu/~purcell/plink/gplink.shtml: 2010.

[9] V. Nitesh , K. W. Chawla, L. O. Bowyer, W. Hall, and K. Philip, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, no. 321, p. 357, 2002.

[10] A. Macchia, G. Levantesi, R. M. Marfisi, M. G. Franzosi, A. P. Maggioni, G. L. Nicolosi, C. Schweiger, L. Tavazzi, G. Tognoni, F. Valagussa, R. Marchioli , and Investigadores del estudio Grupo Italiano para el Estudio de la Supervivencia en el Infarto de Miocardio Prevenzione., "Determinants of Late-Onset Heart Failure in Myocardial Infarction Survivors: GISSI Prevenzione Trial Results," *Rev Esp Cardiol*, vol. 58, pp. 1266-1272, 2005.

[11] B. Kuch, M. Heier, W. Von Scheidt, B. Kling, A. Hoermann, and C. Meisinger, "Meisinger 20-year trends in clinical characteristics, therapy and short-term prognosis in acute myocardial infarction according to presenting electrocardiogram: the MONICA/KORA AMI Registry (1985-2004)," *J Intern Med*, vol. 264, pp. 254-264, 2008.

[12] A. Di Chiara, F. Chiarella, S. Savonitto, D. Lucci, L. Bolognese, S. De Servi, C. Greco, A. Boccanelli, P. Zonzin, S. Coccolini, and A. P. Maggioni, "Epidemiology of acute myocardial infarction in the Italian CCU network: the BLITZ study," *Eur Heart J*, vol. 24, no. 1616, p. 1629, 2003.

[13] G. Montalescot, J. Dallongeville, E. Van Belle, S. Rouanet, C. Baulac, and A. Degrandsart, "STEMI and NSTEMI: are they so different? 1 year outcomes in acute myocardial infarction as defined by the ESC/ACC definition (the OPERA registry)," *Eur Heart J.*, vol. 28, pp. 1409-1417, 2007.

[14] W. J. Rogers, P. D. Frederick, E. Stoehr, J. G. Canto, J. P. Ornato, C. M. Gibson, C. V. Pollack, J. M. Gore, N. Chandra-Strobos, E. D. Peterson, W. J. French, and and for the National Registry of Myocardial Infarction Investigators, "Trends in presenting characteristics and hospital mortality among patients with ST elevation and non-ST elevation myocardial infarction in the National Registry of Myocardial Infarction from 1990 to 2006," *Am Heart J*, vol. 156, no. 6, pp. 1026-1034, 2008.

[15] B. Kuch, R. Wende, M. Barac, W. Von Scheidt, B. Kling, and C. Meisinger, "Prognosis and outcomes of elderly (75-84years) patients with acute myocardial infarction 1-2years after the event - AMI-elderly study of the MONICA/KORA Myocardial Infarction Registry ," *Int J Cardiol*, 2010.

[16] P. Jose, H. Skali, N. Anavekar, C. Tomson, H. M. Krumholz, J. L. Rouleau, L. A. Moye, M. A. Pfeffer , and S. D. Solomon, "Increase in creatinine and cardiovascular risk in patients with systolic dysfunction after myocardial infarction," *J Am Soc Nephrol*, vol. 17, pp. 2886-2891, 2006.

[17] J. Schmitt, G. Duray, B. J. Gersh, and S. H. Hohnloser, "Atrial fibrillation in acute myocardial infarction: a systematic review of the incidence, clinical features and prognostic implications. " *Eur Heart J*, vol. 30, pp. 1038-1045, 2009.

[18] A. Goldberg, H. Hammerman, S. Petchereski, A. Zdorovyak, S. Yalonetsky, M. Kapeliovich, Y. Agmon, W. Markiewicz, and D. Aronson, "Inhospital and 1-year mortality of patients who develop worsening renal function following acute ST-elevation myocardial infarction," *Am Heart J*, vol. 150, pp. 330-337, 2005.

[19] E. Nikolsky, E. Aymong, A. Halkin, C. Grines, D. Cox, E. Garcia, R. Mehran, J. Tcheng, J. Griffin, G. Guagliumi, T. Stuckey, M. Turco, D. Cohen, M. Negoita, A. Lansky, and G. Stone, "Impact of anaemia in patients with acute myocardial infarction undergoing primary percutaneous coronary intervention: analysis from the Controlled Abciximab and Device Investigation to Lower Late Angioplasty Complications (CADILLAC) ," *Trial J Am Coll Cardiol*, vol. 44, pp. 547-553, 2004.

[20] F. Provost and T. Fawcett, "Robust Classification for Imprecise Environments," *Machine Learning*, vol. 42, no. 3, pp. 203-231, 2001.

[21] M. Kubat, R. Holte, and S. Matwin, "Machine Learning for the Detection of Oil Spills in Satellite Radar Images," *Machine Learning*, vol. 30, pp. 195-215, 1998.

[22] C. Ling and C. Li, "Data Mining for Direct Marketing Problems and Solutions," New York, NY: AAAI Press, 1998.

[23] A. Solberg and R. Solberg, "A Large-Scale Evaluation of Features for Automatic Detection of Oil Spills in ERS SAR Images," (Lincoln, NE): 1996, pp. 1484-1486.

[24] P. Domingos, "Metacost: A General Method for Making Classifiers Cost-sensitive," San Diego, CA: ACM Press, 1999, pp. 155-164.

[25] E. DeRouin, E. Brown, L. Fausett, and M. Schneider, "Neural Network Training on Unequally Represented Classes," New York: ASME Press, 1991, pp. 135-141.

[26] D. Lewis and J. Catlett, "Heterogeneous Uncertainity Sampling for Supervised Learning," San Francisco, CA: Morgan Kaufmann, 1994, pp. 148-156.

[27] S. Dumais, J. Platt, D. Heckerman, and M. Sahami, "Inductive Learning Algorithms and Representations for Text Categorization," 1998, pp. 148-155.

[28] D. Mladenic and D. Grobelnik, "Feature Selection for Unbalanced Class Distribution and Naive Bayes," Morgan Kaufmann, 1999, pp. 258-267.

[29] D. Lewis and M. Ringuette, "A Comparison of Two Learning Algorithms for Text Categorization," 1994, pp. 81-93.

[30] W. Cohen, "Learning to Classify English Text with ILP Methods," Department of Computer Science, Katholieke Universiteit Leuven: 1995, pp. 3-24.

[31] T. M. Ha and H. Bunke, "Off-line, Handwritten Numeral Recognition by Perturbation Method," *Pattern Analysis and Machine Intelligence*, vol. 19, no. 5, pp. 535-539, 1997.

[32] N. Japkowicz, " Learning from imbalanced data sets: A comparison of various strategies," 2000.

[33] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," Los Altos, CA : Morgan Kaufmann, 1995, pp. 1137-1143.

[34] M. L. Ginsberg, *Essentials of Artificial Intelligence*. Los ALtos, CA : Morgan Kaufmann, 1993.

[35] P. Norvig and S. J. Russell, *Artificial Intelligence: A Modern Approach*. Englewood Cliffs, NJ : Prentice Hall, 1995.

[36] T. Dean and M. Boddy, "Solving time-dependent planning problems," Detroit, MI: Morgan Kaufmann, 1989, pp. 979-984.

[37] C. Bishop, *Pattern Recognition and Machine Learning*, 1st ed ed. New York: Springer, 2006.

[38] P. N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining* Pearson Education Inc, 2006.

[39] J. Aha, D. Kibler, and M. Albert, "Instance - Based Learning Algorithms," *Machine Learning*, vol. 6, pp. 37-66, 1991.

[40] G. Demiroz and H. Guvenir, "Classification by Voting Feature Intervals," *Lecture Notes In Computer Science*, vol. 1224, pp. 85-92, 1997.

[41] R. Kohavi, "The Power of Decision Tables," Vienna: 1995, pp. 174-189.

[42] M. Hall and E. Frank, "Combining Naive Bayes and Decision Tables," *Association for the Advancement of ArtificialIntelligence*, 2008.

[43] W. Cohen, "Fast effective rule induction," Morgan Kaufmann, 1995, pp. 115-123.

[44] M. Brent, "Instance-Based learning: Nearest Neighbor With Generalization." University of Waikato, Department of Computer Science, 1995.

[45] S. Salzberg, "A nearest hyperrectangle learning method," *Machine Learning*, vol. 6, pp. 277-309, 1991.

[46] E. Frank and I. H. Witten, "Generating Accurate Rule Sets Without Global Optimization," 1998, pp. 144-151.

[47] J. Furnkranz, "Prunning algorithms for rule learning," *Machine Learning*, vol. 27, no. 2, pp. 139-171, 1997.

[48] R. J. Quinlan, *C4.5: programs for machine learning*. San Francisco, CA: Morgan Kaufmann, 1993.

[49] L. Breiman , "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5-32, 2001.

[50] E. E. Tripoliti, D. I. Fotiadis, M. Argyropoulou, and G. Manis, "A six stage approach for the diagnosis of the Alzheimer's disease based on fMRI data," *Journal of Biomedical Informatics*, vol. 43, no. 2, pp. 307-320, 2010.

[51] J. S. Saczynski, D. McManus, J. Yarzebski, D. Lessard, J. M. Gore, and R. J. Goldberg, "Trends in atrial fibrillation complicating acute myocardial infarction," *Am J Cardiol*, vol. 104, pp. 169-174, 2009.