Classifying the Severity of an Acute Coronary Syndrome by Mining Patient Data

Niklas Lavesson* Anders Halling† Michael Freitag‡ Jacob Odeberg§
Håkan Odeberg¶ Paul Davidsson‖

Abstract

An Acute Coronary Syndrome (ACS) is a set of clinical signs and symptoms, interpreted as the result of cardiac ischemia, or abruptly decreased blood flow to the heart muscle. The subtypes of ACS include Unstable Angina (UA) and Myocardial Infarction (MI). Acute MI is the single most common cause of death for both men and women in the developed world. Several data mining studies have analyzed different types of patient data in order to generate models that are able to predict the severity of an ACS. Such models could be used as a basis for choosing an appropriate form of treatment. In most cases, the data is based on electrocardiograms (ECGs). In this preliminary study, we analyze a unique ACS database, featuring 28 *variables, including: chronic conditions, risk factors, and laboratory results as well as classifications into MI and UA. We evaluate different types of feature selection and apply supervised learning algorithms to a subset of the data. The experimental results are promising, indicating that this type of data could indeed be used to generate accurate models for ACS severity prediction.*

*Blekinge Institute of Technology, Box 520, SE-372 25 Ronneby, email: Niklas.Lavesson@bth.se

†Blekinge Competence Center, Vårdskolevägen 5 SE-371 41 Karlskrona, email: Anders.Halling@ltblekinge.se

‡Herlev Hospital, Region Hovedstaden, Herlev Ringvej 75, 2730 Herlev, Denmark

§Dept. of Medicine, Karolinska Institutet and University Hospital, 171 77 Stockholm

¶Blekinge Competence Center, Vårdskolevägen 5, SE-371 41 Karlskrona, email: hakan.odeberg@ltblekinge.se

‖Blekinge Institute of Technology, Box 520, SE-372 25 Ronneby, email: Paul.Davidsson@bth.se

1 Introduction

The ability to identify patients at high risk of morbidity or mortality grows in importance as a consequence of the increasing ability of modern medicine to provide costly but potentially beneficial treatment [3]. Heart disease is the single most common cause of death for both men and women in the developed world [12]. Moreover, it is also one of leading causes of morbidity and mortality in developing countries such as China [4].

When patients with chest pain arrive at the hospital, the physician needs to make an initial diagnosis. However, the consequences of diagnostic errors can be significant for both patients and their physicians [11]. It would therefore be beneficial if the severity of each case could be determined with greater certainty at this initial stage.

The aim of this preliminary study is to investigate the possibility of automatically generating models (classifiers) that can be used to support the diagnosis of Acute Coronary Syndrome (ACS) patients. ACS Patients are difficult to diagnose and they represent a heterogeneous group with different treatment options. Especially for patients presenting to the hosptial early after debut of symptoms and without characteristic electrocardiogram changes of larger myocardial infarction (ST-elevation, see below), no single laboratory marker/test in clinical use today has sufficient diagnostic specificity and sensitivity. Hence, the diagnosis of ACS patients using a data mining approach would be advantageous in many situations [5].

Based on the chronic conditions, risk factors, and laboratory results of a patient, the generated classifier would suggest a diagnosis for that patient. In addition, some types of classifiers are able to motivate their diagnoses by providing rules or trees

that describe the decision process. Unlike opaque models, these transparent classifiers can be used by physicians and other professionals in order to better understand which factors influenced the diagnosis. The decision rules and trees may also contribute to the generation of hypotheses regarding ACS. The outline for the remainder of this paper is as follows. First, we give a more in-depth description of the problem from a medical point of view. This is followed by a review of related work and a presentation of our approach as well as the aims and objectives of this study. We then describe the data mining experiments and follow up with a review of the results. Finally, we draw conclusions and present some pointers to future work.

1.1 Background

An arteriosclerotic plaque, in the context of the heart, is a swelling in artery walls that contain lipids, calcium and connective tissue. Thrombosis is the formation of a clot or thrombus inside a blood vessel, obstructing the flow of blood through the circulatory system. Thrombosis over plaques occurs because of two different mechanisms, one being endothelial erosion, which could lead to a thrombus being adherent to a plaque. The second mechanism is referred to as plaque disruption, or rupture. Thrombosis is a trigger for cardiac ischemia [13]. An ACS is a set of clinical signs and symptoms, interpreted as the result of cardiac ischemia, or abruptly decreased blood flow to the heart muscle. The subtypes of ACS include Unstable Angina (UA), Non-ST Segment Elevation Myocardial Infarction (NSTEMI), and ST Segment Elevation Myocardial Infarction (STEMI).

The Carlscrona Heart Attack Prognosis Study (CHAPS) [6, 9] has recruited patient material for 843 patients with ACS in Karlskrona during 1992-1996. The material includes 494 patients diagnosed with MI and 349 additional patients diagnosed with UA. For each patient, a number of variables concerning chronic conditions, risk factors, and laboratory results were gathered, including: glucose levels, smoking, hypertension, occurrence of hypercholesterolemia. The laboratory results can be available during the initial evaluation of the patients.Also genetic variables are determined exemplified by the common prothrombotic single polymorphism (Glu298Asp) which affects the function of the endothelial Nitric Oxide Enzyme (eNOS) and thereby availability of NO, an important modulator of hemostasis and vascular tone. There is no distinction between NSTEMI and STEMI cases in the CHAPS database. In other words, both of these subtypes are expressed as type MI. An elevation of the ST-segment of the electrocardiogram indicates a severe transmural ischemia in contrast to the ischemia in NSTEMI which only engage the inner part of the myocardium.

1.2 Related Work

The classification or prediction of coronary heart disease has been extensively studied by the machine learning and data mining communities. For example, the diagnosis of MI was featured as a case study when the CART algorithm was first presented [3]. Additionally, the STATLOG project included a heart disease database, containing 13 attributes, in one of the first large-scale comparative studies on machine learning algorithms [7]. A more recent study [1] uses multivariate regression and recursive partitioning analysis to allow the construction of decision rules and of a neural tree for diagnosis. The performance results, as measured with the area under the ROC curve, are quite good. However, the choice of algorithms and their parameter configurations are not described in detail in the paper, which makes it difficult to perform comparisons. On the contrary, another study properly documents four data mining algorithms and their performance on a data set of more than 1,000 patients but fails to describe the data set attributes [4]. In addition, Artificial Neural Network (ANN) ensembles and Logistic Regression models trained on data from 634 patients have been compared in terms of the Area Under the ROC curve (AUC) [5]. The database consisted of electrocardiograms (ECGs) and data that were immediately available at patient presentation. Results indicate that ANNs outperformed Logistic Regression Models. Several studies have also been conducted on the prognosis of patients. For example, one such study [8] investigated the use of ANNs to predict 30 day adverse outcomes from ACS. The setup of variables as featured in the CHAPS database has not been previously studied in data mining research.

2 Method

In this preliminary study we use a quantitative approach to evaluate the suitability of the CHAPS database as a basis for generating ACS prediction models with data mining algorithms. The CHAPS database has been stratified and divided into two separate sets for training/testing and validation, respectively. In this paper, we will focus on the training/testing set in order to determine which types of algorithms are appropriate for the studied problem. The objectives are to compare the default configurations of commonly applied opaque and transparent data mining algorithms and to perform an initial analysis to determine which factors are relevant for accurate classification of ACS patients. The aim is to gain basic knowledge about model generation from the CHAPS database to enable further and more detailed studies on a smaller number of suitable data mining algorithms.

3 Experiment

The experiment is organized as follows. The CHAPS training/testing data set is first converted to the open source ARFF format to allow for analysis with the Weka machine learning workbench [14]. In order to enable the careful scrutiny and repeatability of evaluation results reported, our description of the results is accompanied with all relevant details. Exact parameter specifications are given when the Weka default parameter configuration has not been used. Table 5 includes the complete list of data set attributes along with descriptions as well as possible values (nominal attributes) or the mean and standard deviation (numeric attributes).

3.1 Data Set Analysis

The training/testing data set consists of 422 instances (subjects) classified as either MI (247 instances) or UA (175 instances). In addition to the class attribute, there are 8 nominal attributes and 19 numeric attributes. The nominal attributes are highlighted in Table 1. For each possible attribute value, we have indicated the number of UA and MI cases along with prior probabilities, p. For each value we also give the odds of MI. The attribute and value pairs with the highest odds are marked with bold. The highest odds for MI classification are given by diabetes = yes followed by eNOS = snphomo and smoking = yes. The numeric attributes have been omitted from this part of the analysis since they need to be discretisized for this purpose.

3.2 Initial Performance Evaluation

We first performed an analysis of the complete set of attributes in the training/testing set (422 instances) by comparing the results of 20 data mining algorithms and a baseline algorithm (ZeroR). We used the Weka default configurations for all algorithms except K-nearest Neighbor (IBk) for which we used $k = 10$ to distinguish it from One-nearest neighbor (IB1). Each algorithm was evaluated by averaging the results of 10 runs of 10-fold cross-validation tests with an initial random seed of 1. We recorded results for two quite different evaluation metrics; accuracy (ACC) and the Area Under the ROC curve (AUC). The results, in terms of both ACC and AUC, are presented in Table 2.

The baseline algorithm, ZeroR, generates classifiers consisting of a single rule with zero antecedents and the majority class as the consequent. Thus, they classify all instances as belonging to the MI class. Since $n = 247$ for the MI class and $n = 175$ for the UA class, ZeroR yields an accuracy score of $247/(247 + 175) = 0.59$. The AUC metric was calculated with respect to UA. Thus, UA instances represent the positive cases and MI instances represent the negative cases. Consequently, the True Positives Rate (TPR) depicts the rate of correct UA classifications and the False Positives Rate (FPR) depicts the rate of MI cases classified as UA. With regard to AUC, the baseline behaves as a random guesser, thus it yields an AUC score of 0.50.

The best AUC score was achieved by the Logistic algorithm (0.74) followed by AdaBoostM1 and Bagging (0.73) while Support Vector Machines (SMO) achieved the best ACC score (0.70) followed by Logistic and Bagging (0.69). When averaging across the two metrics, the overall best performing algorithms were: Logistic and Bagging (0.71), followed by AdaBoostM1 and BayesNet (0.70), and SMO (0.69).

Table 1: Nominal attribute statistics

Attribute	Values	Classification				total	p	MI
		MI	p	UA	p			odds
sex	male	178	0.72	118	0.67	296	0.70	**1.07**
	female	69	0.28	57	0.33	126	0.30	0.86
	missing	0	0.00	0	0.00	0	0.00	
hypertension[a]	no	175	0.71	126	0.72	301	0.71	**0.98**
	yes	64	0.26	47	0.27	111	0.26	0.96
	missing	8	0.03	2	0.01	10	0.02	
diabetes	no	180	0.73	144	0.82	324	0.77	0.89
	yes	52	0.21	20	0.11	72	0.17	**1.84**
	missing	15	0.06	11	0.06	26	0.06	
heart_failure[b]	no	211	0.85	142	0.81	353	0.84	**1.05**
	yes	28	0.11	31	0.18	59	0.14	0.64
	missing	8	0.03	2	0.01	10	0.02	
diabetes_treatment	no	202	0.82	152	0.87	354	0.84	0.94
	pills	7	0.03	0	0.00	7	0.02	0.00
	insulin	4	0.02	4	0.02	8	0.02	0.71
	diet	26	0.11	17	0.10	43	0.10	**1.08**
	missing	8	0.03	2	0.01	10	0.02	
smoking	no	172	0.70	144	0.82	316	0.75	0.85
	yes	63	0.26	28	0.16	91	0.22	**1.59**
	missing	12	0.05	3	0.02	15	0.04	
hypercholesterolemia	no	231	0.94	155	0.89	386	0.91	**1.06**
	yes	8	0.03	18	0.10	26	0.06	0.31
	missing	8	0.03	2	0.01	10	0.02	
eNOS	wildhomo	114	0.46	95	0.54	209	0.50	0.85
	hetero	107	0.43	69	0.39	176	0.42	1.10
	snphomo	26	0.11	11	0.06	37	0.09	**1.67**
	missing	0	0.00	0	0.00	0	0.00	

[a] Treated for high blood pressure
[b] Treated for dysfunction of the heart muscle pump

Table 2: Initial results on the complete set of attributes

Algorithm	Type	AUC	ACC
AdaBoostM1	opaque	0.73(0.07)	0.68(0.07)
Bagging	opaque	0.73(0.07)	0.69(0.06)
BayesNet	opaque	0.72(0.07)	0.67(0.07)
Dagging	opaque	0.69(0.09)	0.65(0.06)
DecisionStump	opaque	0.68(0.06)	0.67(0.07)
HyperPipes	opaque	0.54(0.06)	0.58(0.03)
IB1	opaque	0.54(0.08)	0.54(0.08)
IBk ($k = 10$)	opaque	0.59(0.09)	0.57(0.08)
Logistic	opaque	0.74(0.07)	0.69(0.07)
MLP[a]	opaque	0.65(0.09)	0.62(0.08)
NaiveBayes	opaque	0.69(0.07)	0.58(0.06)
RandomForest	opaque	0.67(0.09)	0.64(0.08)
RBFNetwork	opaque	0.67(0.08)	0.63(0.07)
SMO	opaque	0.68(0.07)	0.70(0.07)
BFTree	transparent	0.67(0.09)	0.68(0.07)
J48	transparent	0.61(0.08)	0.63(0.07)
JRip	transparent	0.65(0.07)	0.66(0.07)
PART	transparent	0.61(0.09)	0.61(0.07)
Ridor	transparent	0.63(0.07)	0.65(0.07)
SimpleCart	transparent	0.69(0.07)	0.67(0.07)
ZeroR	transparent	0.50(0.00)	0.59(0.01)

[a] MultiLayerPerceptron

Table 3: Results on three feature selected data sets

Algorithm	Best First		Nominal		Numeric	
	AUC	ACC	AUC	ACC	AUC	ACC
AdaBoostM1	0.70(0.07)	0.66(0.07)	0.61(0.08)	0.61(0.06)	0.70(0.07)	0.66(0.07)
Bagging	0.75(0.07)	0.62(0.07)	0.62(0.08)	0.61(0.06)	0.68(0.07)	0.58(0.06)
BayesNet	0.76(0.07)	0.71(0.06)	0.61(0.08)	0.60(0.06)	0.75(0.07)	0.70(0.06)
BFTree	0.71(0.08)	0.68(0.07)	0.55(0.09)	0.55(0.06)	0.67(0.08)	0.63(0.06)
Dagging	0.73(0.07)	0.69(0.06)	0.60(0.08)	0.60(0.06)	0.65(0.08)	0.62(0.07)
DecisionStump	0.69(0.06)	0.71(0.06)	0.52(0.04)	0.59(0.05)	0.70(0.06)	0.71(0.06)
HyperPipes	0.72(0.07)	0.70(0.06)	0.57(0.09)	0.56(0.06)	0.66(0.07)	0.64(0.06)
IB1	0.60(0.07)	0.61(0.07)	0.49(0.07)	0.51(0.07)	0.58(0.07)	0.59(0.07)
IBk (k=10)	0.74(0.07)	0.68(0.06)	0.57(0.07)	0.60(0.03)	0.73(0.07)	0.68(0.07)
J48	0.74(0.06)	0.70(0.06)	0.59(0.09)	0.58(0.05)	0.72(0.07)	0.69(0.06)
JRip	0.69(0.08)	0.62(0.05)	0.57(0.10)	0.60(0.04)	0.70(0.08)	0.65(0.06)
Logistic	0.53(0.04)	0.59(0.02)	0.51(0.01)	0.59(0.01)	0.53(0.06)	0.58(0.03)
MLP[a]	0.67(0.08)	0.67(0.07)	0.50(0.06)	0.58(0.04)	0.67(0.09)	0.67(0.07)
NaiveBayes	0.68(0.06)	0.67(0.07)	0.51(0.05)	0.59(0.02)	0.68(0.06)	0.67(0.07)
PART	0.70(0.08)	0.68(0.07)	0.53(0.07)	0.57(0.04)	0.66(0.08)	0.65(0.07)
RandomForest	0.70(0.08)	0.66(0.06)	0.52(0.09)	0.54(0.07)	0.67(0.09)	0.64(0.07)
RBFNetwork	0.69(0.07)	0.68(0.07)	0.51(0.04)	0.57(0.04)	0.69(0.07)	0.67(0.07)
Ridor	0.66(0.07)	0.68(0.07)	0.51(0.06)	0.56(0.06)	0.65(0.08)	0.66(0.07)
SimpleCart	0.68(0.07)	0.65(0.07)	0.57(0.09)	0.57(0.06)	0.68(0.08)	0.65(0.08)
SMO	0.65(0.06)	0.67(0.06)	0.51(0.04)	0.58(0.04)	0.63(0.07)	0.65(0.06)
ZeroR	0.50(0.00)	0.59(0.01)	0.50(0.00)	0.59(0.01)	0.50(0.00)	0.59(0.01)
Average	0.69(0.07)	0.67(0.06)	0.55(0.07)	0.58(0.05)	0.67(0.07)	0.65(0.06)

[a] MultiLayerPerceptron

3.3 Feature Selection

We generated three new data sets using feature selection. The first data set was generated using only numeric attributes (except for the class attribute) while the second data set only features nominal attributes. The third data set, which features 5 numeric and 1 nominal attributes, was generated using the Best First feature selection algorithm [15]. The Best First method is a heuristic search strategy that uses hill climbing and a back-tracking mechanism to reduce the number of attributes and increase the performance [14]. Out of the complete set of attributes, the Best First method selected the following attributes: heart_failure, B_LPK, H1_NEU, B_GLU, B_TMCV, and P_APTT. We again evaluated each algorithm using 10 runs of 10-fold cross-validation tests. The results, which can be viewed in Table 3, indicate that the Best First feature selected data set is the most suitable, since the average AUC and ACC are the highest in comparison to the other data sets, including the data set with the complete set of attributes. BayesNet achieves the highest AUC and ACC, followed by Bagging, J48 and IBk (AUC) and DecisionStump (ACC). Interestingly, Logistic performs poorly on the Best First data set.

3.4 Classifier Understandability

There is often a trade-off between classification performance and understandability. In our experiment, we evaluated several rule and tree based algorithms that are able to produce classifiers that may provide human-understandable visualizations of the classification process. However, the understandability of tree- and rule-based models depends on the complexity of the trees and rule sets. Other models, e.g., generated by SMO, can also be understood in the sense that it can be determined which attributes are important indicators for a particular class. However, related work often seem to treat neural network and support vector machine models as being opaque. As a result, several studies have presented approaches to generate understandable rules from such models, cf. [2]. We provide some rule-based examples in Table 4 and one decision tree can be viewed in Figure 1.

4 Discussion

We used two different evaluation metrics for this purpose. Firstly, we measured the classification accuracy, i.e., the ratio of correctly classified instances. This metric has been the traditional choice for evaluation and it is very straight-forward to use

Table 4: Rule-based classifiers

Algorithm	Classifier
The following rules were produced using the complete set of attributes	
Jrip	IF (**B_LPK** $\leq$ 7.44) THEN **diagnosis** = ua ELSE **diagnosis** = mi

Ridor	IF (**B_LPK** $>$ 8.175) AND (**B_GLU** $>$ 5.935) AND (**P_PT** $\leq$ 101.5)
	AND (**B_MCV** $>$ 82.5)
	THEN **diagnosis** = mi
	IF (**B_LPK** $>$ 8.175) THEN **diagnosis** = mi
	IF (**B_LPK** $>$ 6.345) AND (**B_TMCV** $\leq$ 8.45) AND (**B_GLU** $>$ 5.45)
	THEN **diagnosis** = mi
	ELSE **diagnosis** = ua

ConjunctiveRule	IF (**B_LPK** $>$ 8.095) THEN **diagnosis** = mi
The following rules and the tree in Figure 1 were produced using the Best first selected attributes	

JRip	IF (**B_LPK** $\leq$ 8.2) AND (**B_TMCV** $\geq$ 9) THEN **diagnosis** = ua
	IF (**B_LPK** $\leq$ 7.41) THEN **diagnosis** = ua
	ELSE **diagnosis** = mi

PART	IF (**B_LPK** $>$ 8.09) AND (**B_GLU** $>$ 5.92) AND (**heart_failure** = no)
	THEN **diagnosis** = mi
	IF (**B_LPK** $>$ 8.81) THEN **diagnosis** = mi
	IF (**B_GLU** $\leq$ 4.6) THEN **diagnosis** = ua
	IF (**B_TMCV** $\leq$ 8.5) AND (**H1_NEU** $>$ 4.18) THEN **diagnosis** = mi
	ELSE **diagnosis** = ua

as well as being easily explainable. However, it suffers from the assumption that the class distribution is known for the target domain and it also assumes equal misclassification costs [10]. These two assumptions are rarely met in real-world problems and the studied problem is a perfect example of this. Thus, we also calculated the area under the ROC (AUC) metric for the purpose of classifier evaluation. This metric does not suffer from the two earlier mentioned assumptions; however, it does suffer from an information loss in comparison to a complete ROC plot. To summarize, there are known issues with most of the currently used evaluation metrics, but we argue that the combined information gained from the ACC and AUC evaluations is adequate for the purpose of this preliminary study. We first generated classifiers using the complete set of attributes. The best performing classifiers achieved an accuracy score of 0.70 and an AUC score of 0.74, while the worst performing classifiers behaved like random guessers. These results may be satisfactory for real-world diagnosis purposes; however, we assumed that the results could be improved by reducing the dimensionality of the data set in terms of the number of input attributes. We therefore proceeded by applying a feature selection algorithm to reduce the number of attributes. We used the Best First feature selection method and succeeded in reducing the number of attributes from 27 to 6 while increasing both ACC and AUC for most algorithms. However, the increase in performance was only slight. The best performing classifier now achieved an accuracy score of 0.71 and an AUC score of 0.76.

5 Conclusions

This preliminary study has investigated the potential for using data mining methods to find useful patterns in an Acute Coronary Syndrome (ACS) patient data set. If found, such patterns could be used to generate classifiers that would aid the diagnosis of future ACS subjects. We have trained and evaluated 20 well-known data mining algorithms on different variations of a set of 422 instances. Each instance describes a patient by using 27 input attributes, diagnosed as either Unstable Angina ($n = 175$) or Myocardial Infarction ($n = 247$). The performance results are promising; however, we speculate that the access to more training data and careful parameter tuning could increase the performance further. This study also shows that the featured opaque classifiers perform better than the transparent (understandable) classifiers. This makes it interesting to further explore the trade-off between classification performance and understandability. However, one notable exception to this rule is the J48 tree inducer, which managed to achieve an AUC score of 0.74 on the Best First

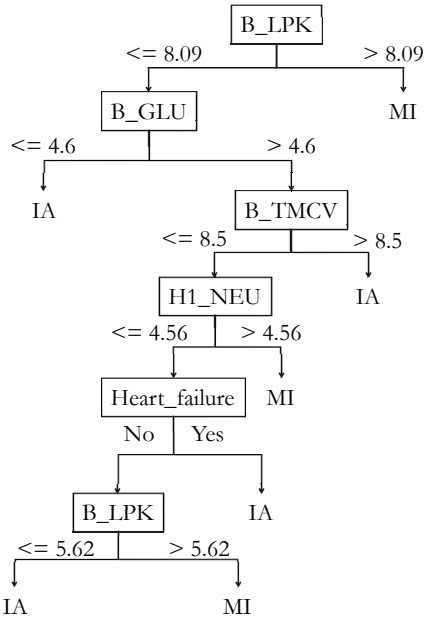

Figure 1: J48 decision tree with 6 branches and 7 leaves

data set. Perhaps most interestingly, most learning algorithms, as well as the feature selection algorithm, tended to agree on the importance of at least two attributes: B_LPK and B_GLU. For example, JRip managed to achieve an accuracy of 0.66 by generating a rule based only on B_LPK. There are a number of interesting directions for future work. Firstly, we would like to establish which feature selection method is the most suitable for the domain. We also intend to perform extensive algorithm parameter tuning in order to generate better models by concentrating on the best performing algorithms from this study. The aim is to validate the results of these new models by perform evaluations on the previously unseen validation data set. Thirdly, we will perform a deeper analysis of the featured attributes and investigate correlations between them. We would also like to introduce additional attributes describing inflammatory markers that may be suitable indicators of the severity of an ACS outcome.

Acknowledgments

This work was partly funded by Blekinge County Council.

References

[1] R. Bassan, L. Pimenta, M. Scofano, and J. F. Soares. Accuracy of a neural diagnostic tree for the identification of acute coronary syndrome in patients with chest pain and no st-segment elevation. *Critical Pathways in Cardiology*, 3(2):72–78, 2004.

[2] Ricardo Blanco-Vega, José Hernández-Orallo, and M. José Ramírez-Quintana. Analysing the trade-off between comprehensibility and accuracy in mimetic models. In *Discovery Science*, pages 338–346, 2004.

[3] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. *Classification and Regression Trees*. 1984.

[4] J. Chen, Y. Xing, G. Xi, J. Chen, J. Yi, D. Zhao, and J. Wang. A comparison of four data mining models: Bayes, neural network, svm and decision trees in identifying syndromes in coronary heart disease. In *Fourth International Symposium on Neural Networks*, 2007.

[5] M. Green, J. Björk, J. Forberg, U. Ekelund, L. Edenbrandt, and M. Ohlsson. Comparison between neural networks and multiple logistic regression to predict acute coronary syndrome in the emergency room. *Artificial Intelligence in Medicine*, 38(3):305–318, 2006.

[6] K. Holmberg, M-L. Persson, M. Uhlén, and J. Odeberg. Pyrosequencing analysis of thrombosis-associated risk markers. *Clinical Chemistry*, 51:1549–1552, 2005.

[7] R. D. King, C. Feng, and A. Sutherland. STATLOG: Comparison of classification algorithms on large real-world problems. *Applied Artificial Intelligence*, 9(3):259–287, 1995.

[8] C. L. McCullough, A. J. Novobilski, and F. M. Fesmire. Use of neural networks to predict adverse outcomes from acute coronary syndrome

for male and female patients. In *Sixth International Conference on Machine Learning and Applications*, 2004.

[9] J. Odeberg, M. Freitag, H. Odeberg, L. Råstam, and U. Lindblad. Severity of acute coronary syndrome is predicted by interactions between fibrinogen concentrations and polymorphisms in the GPIIIa and FXIII genes. *Thrombosis and Haemostasis*, 4:909–912, 2006.

[10] Foster Provost, Tom Fawcett, and Ron Kohavi. The case against accuracy estimation for comparing induction algorithms. In *15th International Conference on Machine Learning*, pages 445–453, San Francisco, CA, USA, 1998. Morgan Kaufmann Publishers.

[11] R. A. Rusnak, T. O. Stair, K. Hansen, and J. S. Fastow. Litigation against the emergency physician: Common features in cases of missed myocardial infarction. *Annals of Emergency Medicine*, 18:1029–1034, 1989.

[12] H. Tunstall-Pedoe, K. Kuulasmaa, P. Amouyel, D. Arveiler, A. M. Rajakangas, and A. Pajak. Myocardial infarction and coronary deaths in the World Health Organization MONICA project. registration procedures, event rates, and case-fatality rates in 38 populations from 21 countries in four continents. *Circulation*, 90(1):583–612, 1994.

[13] L. Wallentin and B. Lindahl A. Siegbahn. Unstable coronary artery disease. In E. Falk, editor, *Textbook of Cardiac Disease*. Mosby, New York, 2002.

[14] Ian H. Witten and Eibe Frank. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann Publishers, San Francisco, CA, USA, 2005.

[15] L. Xu, P. Yan, and T. Chang. Best first strategy for feature selection. In *Ninth International Conference on Pattern Recognition*, pages 706–708, New york City, NY, USA, 1988. IEEE Press.

Table 5: Data Set Description

Attribute	Values[a]	Description[b]
sex	male,female	
age	63.8(8.78)	
hypertension	no,yes	
diabetes	no,yes	
heart_failure	no,yes	
diabetes_treatment	no,pills,insulin, diet	
smoking	no,yes	
hypercholesterolemia	no,yes	
eNOS	wildhomo[c],hetero,[d] snphomo[e]	endothelial Nitric Oxide Synthase
B_LPK	8.94(3.03)	B-Leucocytes
B_HB	136.7(14.5)	B-Hemoglobin
B_EVF	40.5(4.24)	B-Hematocrit
B_MCV	90.37(5.43)	B-Erytrocyte Mean Corpuscular Volume
B_TROM	226.1(63.7)	B-Thrombocytes
H1_NEU	6.27(2.78)	B-Neutrophils
P_PT	83.97(24.49)	P-Prothrombin Time
S_KREA	101.9(71.43)	S-Creatinine
S_ALB	38.33(3.65)	S-Albumin
S_KOL	6.17(1.32)	S-Cholesterol
S_HDLKOL	1.17(0.38)	S-HDL-Cholesterol
B_GLU	6.78(3.05)	B-Glucose
S_TSH	2.18(3.04)	S-Thyroid-Stimulating Hormone
B_TMCV	9.04(0.75)	B-Thrombocyte Mean Corpuscular Volume
P_APTT	33.17(21.60)	P-Activated Partial Thromboplastin Time
S_TG	2.03(1.43)	S-Triglycerides
S_HBA1C	5.24(1.31)	S-Hemoglobin A_{1C}
P_FGEN	3.69(0.94)	P-Fibrinogen
diagnosis	mi,ua	

[a] Given as the complete set of categories (nominal) or the mean and SD (numeric)
[b] The laboratory samples are of type: Blood (B), Serum (S), or Plasma (P)
[c] Wild-type homogeneous eNOS
[d] Heterogeneous eNOS
[e] Single-nucleotide polymorphism eNOS