

# Real-time Decision Support using Data Mining to predict Blood Pressure Critical Events in Intensive Medicine Patients

Filipe Portela<sup>1</sup>, Manuel Filipe Santos<sup>1</sup>, José Machado<sup>2</sup>, António Abelha<sup>2</sup>, Fernando Rua<sup>3</sup> and Álvaro Silva<sup>3</sup>

<sup>1,2</sup> Algoritmi Centre, University of Minho, Portugal;

<sup>1</sup> [cfp@dsi.uminho.pt](mailto:cfp@dsi.uminho.pt), [mfs@dsi.uminho.pt](mailto:mfs@dsi.uminho.pt); <sup>2</sup> [jmac@di.uminho.pt](mailto:jmac@di.uminho.pt), [abelha@di.uminho.pt](mailto:abelha@di.uminho.pt)

<sup>3</sup> Serviço Cuidados Intensivos, Centro Hospitalar do Porto, Hospital Santo António

<sup>3</sup> [moreirasilva@me.com](mailto:moreirasilva@me.com), [fernandorua.sci@chporto.min-saude.pt](mailto:fernandorua.sci@chporto.min-saude.pt)

**Abstract.** Patient blood pressure is an important vital signal to the physicians take a decision and to better understand the patient condition. In Intensive Care Units is possible monitoring the blood pressure due the fact of the patient being in continuous monitoring through bedside monitors and the use of sensors. The intensivist only have access to vital signs values when they look to the monitor or consult the values hourly collected. Most important is the sequence of the values collected, i.e., a set of highest or lowest values can signify a critical event and bring future complications to a patient as is Hypotension or Hypertension. This complications can leverage a set of dangerous diseases and side-effects. The main goal of this work is to predict the probability of a patient has a blood pressure critical event in the next hours by combining a set of patient data collected in real-time and using Data Mining classification techniques. As output the models indicate the probability (%) of a patient has a Blood Pressure Critical Event in the next hour. The achieved results showed to be very promising, presenting sensitivity around of 95%.

**Keywords:** Data Mining, INTCare, Intensive Medicine, Blood Pressure, Critical Events, Decision Support, Real-Time

## 1 Introduction

In critical environments the decision needs to be perform quickly and with a high level of accuracy. To help the decision-makers to take the best decision it is fundamental to develop a solution able to predict events before their occurrence. Intensive Medicine (IM) is a critical area of Medicine. Patients in weak conditions and with multiple diseases as is organ failure are cared every day. One of the most common complications is related to Blood Pressure with a constant values changing due to medical diseases, therapeutics or other procedures. Higher Blood pressure is associated to cardiovascular organ failure / diseases [1]. Nowadays the Intensive Care Units (ICU) are filled out with many technical devices allowing a continuous patient monitoring. However these data are only used in the acquisition moment and they are not used to support the decision process. Having in consideration this aspect arises INTCare. INTCare [2] is a Pervasive

Intelligent Decision Support System (PIDSS) able to collect and process data in real-time in order to provide new knowledge [3-7] anywhere and anytime. This knowledge is achieved by means of Data Mining (DM) techniques. This work is framed in INTCare project and it wants to develop DM models able to help the Intensivist to act in order to prevent Blood Pressure Critical Events (BPCE). Critical Events are defined as a continuous data acquisition of values out of the normal range for a determined period of time. BPCE can provoke hypotension and hypertension and leverage a set of other diseases as is heart attack, cardiovascular system failure, kidney failure and others. This work is a classification DM problem with the goal to induce DM models in real-time able to predict the probability of a patient has a BPCE in the next hours. The achieved results are promising. The models are very good to predict BPCE (sensitivity around 95%) however they are not accurate in predicting both classes (accuracy around 75%). These results are a natural consequence of preventing actions in IM. The predictions are made in an hourly base approach and due to a quick decision by the intensivist, the BP value can change to a normal range in a few minutes resulting in a set of false positives values. In a clinical point of view, these models are very useful because can help to predict a possible CE in the next hours giving the possibility to the intensivist take a decision based on evidences i.e., according to the new knowledge achieved. Supporting Decision-Making process by providing the probability of a patient has a critical event in the next hours in order to prevent and avoid their occurrence is the main goal of this work. The models developed will be included in INTCare system.

This paper is divided in six sections. After a first introduction of the paper, all the concepts and related work are presented in section 2. Section 3 presents the methodologies, materials and methods used in this work. Section four presents all the work made following CRISP-DM phases. Then the results are discussed and a set of last considerations are made including future work.

## **2 Background**

### **2.1 Blood pressure**

According to the National Institute of Health [8], Blood Pressure (BP) is the force of blood pushing against the walls of the arteries as the heart pumps blood. If this pressure rises and stays high over time, it can damage the body in many ways. BP is measured as systolic and diastolic pressures. "Systolic" refers to BP when the heart beats while pumping blood. "Diastolic" refers to BP when the heart is at rest between beats. BP has a normal pattern of values. Values out of this pattern can provide Hypotension (low pressure) and Hypertension (Higher pressure). In clinical studies [9] was found a strong relationship between prevalence of hypertension and mortality by stroke. Hypertension can provoke several and dangerous diseases to human life [10] as is for example: Heart attack or stroke, Aneurysm Heart failure and others. Hypotension is dangerous and can be a source of other diseases [11] as is for example Kidney Failure, Congestive Heart Failure, Anaemia, and Pulmonary Embolism.

In the literature review there are several works related to BP [12], however the goals are different. This work predicts the probability of occurring a BPCE in the next hours

(e.g. Hour1-10%...Hour4-15%...Hour24-30%) by using streaming data and intelligent agents to perform automatically tasks (e.g. data transformation and induce the models.)

## 2.2 Intensive Care Units

Intensive Medicine (IM) is a critical area with the highest incidences of medical error and patient injuries [13]. Its practices is in Intensive Care Units (ICUs). To the ICUs only patients in a critical condition are transferred. ICUs are endowed with several medical devices (e.g. vital signs monitor, ventilators and infusion pumps) able to monitoring the patient condition in real-time. Additionally a set of clinical data are hourly recorded. Although there are a high number of data available, the information are not fully used to create new knowledge. This fact happens due to the need of having a quick decision in order to avoid worst conditions and save patient life. In this sense the intensivists do not have time to consult or analyse all the collected data. Thus, becomes fundamental having a decision support system able to help them to take the best decision always in the patient best interest by considering all the collected data.

## 2.3 INTCare

INTCare is a PIDSS able to monitoring the patient condition in real-time and to predict a set of clinical events/diseases using the collected data and adaptive DM models. Online-learning, real-time data processing [14] and system interoperability are others features. Until now, INTCare allows predicting organ failure and patient outcome [7], SEPSIS [15], barotrauma [16], readmissions [5, 17] and length of stay [6, 18] in real-time and with high sensitivities rates. In the past a first study was performed in order to predict critical events. In this study [4] the probability of a patient having a cardiac arrhythmia was predicted (sensitivity = 95%). Now it is time to explore other areas and tracking the Blood Pressure in order to early detect possible critical events, i.e., possible situations of hypertension or hypotension.

## 2.4 Critical Events

Critical Events (CE) represents an abnormal value (out of normal range) continuously monitored in a patient. A CE is defined by Álvaro Silva [19] as “a more serious event and it is classified by a longer event or a more extreme out of range measurement”. In the ICU the vital signs are collected and processed in real-time [20] in order to categorize a value as critic or not. Then an intelligent agent is used to analyse if the values collected can represent or not a critical event. To understand if an event is critical, two main criteria were used [21]:

- Occurrence and duration should be registered by physiological changes;
- Related physiological variables should be registered at regular intervals.

An event is considered critical, when a longer event occurs or a more extreme physiological measurement is found [21]. In this work the protocol associated to Blood Pressure (Table 1) was followed. For example, a critical event happens whenever the patient's blood pressure is less than 90 mmHg for more than 1 hour or the value drops below 60 mmHg.

**Table 1.** The protocol for the out of range physiologic measurements (adapted from [21] )

	<b>Blood Pressure</b>
Normal range	90—180 mmHg
Critical event (continuously out of the normal range)	>= 1h
Critical event (anytime)	< 60 mmHg

### 3 Methodologies, Material and Methods

In this work the Design Science Research Methodology (DSR) was followed. DSR is fundamental in developing effective solutions - products, services, and systems able to answer to human needs [22]. According to Lunenfeld [23] research for design is the hardest to characterize, its purpose is to create objects and systems that display the results of the research and prove its worth. DSR is based in the creation and assessment of artefacts. To complement this methodology, Cross Industry Standard Process for Data Mining (CRISP-DM) was used. It is divided in six phases: Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation and Deployment. By crossing DSR and CRISP-DM it was possible develop and assess an artefact (prediction models) able to support the decision-making in the BP field.

This study used the data collected in real-time from 359 patients admitted in the ICU of Hospital Santo António, Centro Hospitalar do Porto, comprising a period between 2012.02.01 to 2014.02.26 (757 days) in a total of 222381 rows. In this project four data systems were considered: Vital Signs Monitors, Laboratory, Electronic Health Record and Pharmacy. Data mining models were induced exploring four different techniques: Decision Trees, Support Vector Machines, Naïve Byes and Generalized Linear Models.

### 4 Real-Time Decision Support

As already mentioned all the work was developed recurring to CRISP-DM methodology. The main work developed in each phase is presented in this section.

#### 4.1 Business Understanding

The problem and their importance to the environment (service, professionals and patients) was already presented in the section 2.1. In this work to avoid the occurrence of hypertension and hypotension in patients admitted in ICU is the business goal. Developing models with a high level of sensitivity able to support the decision process by predicting the probability of a patient has a BPCV in the next hours is the DM goal.

#### 4.2 Data Understanding

In this phase the initial dataset was analysed in order to prepare it to be used by the DM engine. With the goal to answer to the main question it was used data provided from four data sources: Vital Signs Monitors – {Blood Pressure and Heart Rate}; Electronic Health

Records – {Admission variables and Age}; Therapeutic Plan – {Vasopressores}; Laboratory Results – {Bilirubin, Creatinine, Po2/Fio2 and Platelets}.

In table 2 is presented a distribution (percentage of cases) of the non-numeric variables, age and target (20% of the records has critical events associated). These variables were not submitted to a processing phase because the value was only associated to a pre-defined class. It represents a simple matching between the DM classes and the values collected for each variable.

**Table 2.** Variables Distribution

<i>ID</i>	<i>Variable</i>	<i>Min</i>	<i>Max</i>	<i>Class</i>	<i>Cases</i>
Age	Age	18	46	1	17.37%
		47	65	2	35.46%
		66	75	3	21.32%
		76	130	4	25.85%
Admission Type	Urgent	-	-	U	80.29%
	Programmed	-	-	P	19.71%
Admission Provenance	Chirurgic	-	-	1	47.66%
	Observation	-	-	2	0.05%
	Emergency	-	-	3	18.44%
	Nursing Room	-	-	4	15.37%
	Other ICU	-	-	5	2.44%
	Other Hospital	-	-	6	1.32%
	Other	-	-	7	14.73%
Insufficiencies Cardiac	Yes	-	-	1	91.76%
Transplant	Yes	-	-	1	10.24%
Surgical admission	Yes	-	-	1	55.02%
Cerebrovascular Accident (CVA)	Yes	-	-	1	2.20%
Critical Event	Yes	-	-	1	20.00%

### 4.3 Data Preparation

In this phase all the variables were validated. First the existence of null values was verified and then it was verified the occurrence of values out of the acceptable range (noise values). Then all transformations rules were executed. Both tasks were performed by intelligent agents. The tasks performed were:

- To verify and group the admission type and admission from;
- To classify the Blood Pressure values as critical or not;
- To create the critical events variable;
- To determine the patient SOFA value for all the organic systems;
- To create a new variable which identifies if the patient has risk factors;
- To determine the last seven Blood Pressure values (BPLV) collected;
- To calculate the accumulated critical events (ACE) and all associated ratios;
- To create classes to numerical values (e.g. ratios and BPLV).
- To create the DM input dataset in a hourly base;

To induce the DM models, several scenarios were prepared using attributes sets. The first set created was Case Mix (CM). All the CM variables were provided by the patient Electronic Health Record (EHR). From the EHR, the variables: age, admission type, admission from, insufficiencies cardiac and risk patient (combination of a set of patient admission variables) were used. These variables were recorded at patient admission phase, being then transformed in accordance with the DM attributes.

The second group of variables used was the Sequential Organ Failure Assessment (SOFA). SOFA is used in ICU to score the degree of dysfunction/failure of the six organic systems (cardiovascular, respiratory, renal, liver, haematological and neurological) [24]. SOFA score varies from 0 (normal function) and 4 (total dysfunction). In this case the transformation made was simple: in case of normal function it was attributed the value 0 otherwise the value was 1. Thus, the attribute SOFA used in DM only considers two values 0 (SOFA = 0) or 1 (SOFA > 0).

The third group was based in the CE concept by creating the Accumulated Critical Events (ACE) variable. ACE includes three physiological variables: Blood Pressure (BP), Saturation of Oxygen (SPO2) and Heart Rate (HR). ACE values are calculated in real-time by counting the number of critical events verified by hour since patient admission. An attribute (TOTAL) that reflects the sum of ACE by each variable was also added. Then and using ACE a set of ratios were introduced.

These ratios allow to determine a relation between the number of ACE verified in a patient and the maximum number of ACE occurred by hour (R1) and a correspondence between the number of ACE verified in a patient and the maximum number of events verified in the past (from all the patients), until the hour in analysis (R2). Both the values were grouped by category and by patient.

Finally another group composed by the last seven BP values (BPLV) collected in a specific hour was created. This attribute is generated all hours by an agent. It analyse the last values collected in the last hour and put them in the same row (hour).

After performing all the transforming tasks, a set of new variables were introduced. These variables and their distribution are presented in the table 3. For example, the SOFA\_HEPATIC attribute is 1 when the bilirubin value collected is higher than the minimum considered (1.2). In the case of Risk Patient, the attribute is 1 when at least one of the symptoms or conditions presented in the table 3 is verified. These conditions were verified in 21.32% of the cases.

**Table 3.** Transformed variables distribution

<i>Attribute</i>	<i>Variable</i>	<i>Min</i>	<i>Max</i>	<i>Value</i>	<i>% Cases</i>
SOFA_CARDIO	BP (mean)	0	70	1	72.63%
	Dopamine or Dobutamine	0.0	-	1	
	Epi / Norepi	0.0	-	1	
SOFA_RENAL	Creatinine	1.2	-	1	21.41%
SOFA_RESPIRAT	Po2/Fio2	0	400	1	67.76%
SOFA_HEPATIC	Bilirubin	1.2	-	1	21.37%
SOFA_HEMAT	Platelets	0	150	1	47.76%
RISK PATIENT	CVA	-	-	1	21.32%
	Alcoholism or Addicted	-	-	1	
	Pacemaker	-	-	1	
	Corticoids	-	-	1	
	Transplanted	-	-	1	
	Vasoactive Drug	-	-	1	

In the last step a discretization technique has been considered. Having in consideration all numeric values used by the models, the attributes were transformed and categorized according to an interval (*Min* and *Max*). Using this technique the groups were defined according to the respective average (R1) or higher value (R2) of the data collected. These ranges are flexible and they are updated automatically according to the

values collected. The groups were defined using a 7-point-scale adapted by Clinical Global Impression - Severity scale (CGI-S) [25]. The goal of CGI-S is to allow the clinician to rate the severity of illness [26]. The ranges defined [7] has clinical significance and represents the patient condition. A higher value represents a worst patient condition. The standard used to define the percentages concentrate the most part of patient values within a scale between 0 and 5. More severe cases are assigned to the levels 6 and 7. Table 4 presents the rules defined to discretize each continuous value ( $values \in \{R_{0+}\}$ ). The groups are identified at the top of the table and the left column identifies the variable. In the centre of the table are the ranges for each group. These values were obtained after apply the ranges (%) defined.

The attributes of  $R1$  are determined by the rows ( $R1$  BP Min to  $R1$  TOT Max).  $R2$  is categorized according to the percentage of the values collected, i.e., level 1 corresponds to 10% of the cases (values collected between 0% and 10% of the Maximum). All attributes of  $R2$  have the same limits. ACE attributes were grouped in agreement to their importance and the number of occurrences. These values were defined by ICU experts but can be modified in the future.

In case of ACE, if a BP value equal to 2 is verified, this variables is categorized in the second level (value = 1). The same happens to patients who present values of  $R2$  between 0.01 and 0.10 or  $R1$  HR between 0.004 and 0.008. The CGI-S was not used to calculate BPLV. The clinical guidelines and the values defined as normal were considered. Level 2 and 3 corresponds to the acceptable (normal) range.

**Table 4.** Discretization sets of Data Mining Input

SET		0	1	2	3	4	5	6	7
R1	Min	-0.1	0.000	0.010	0.021	0.041	0.062	0.082	0.123
BP	Max	0.000	0.010	0.021	0.041	0.062	0.082	0.123	2.000
R1	Min	-0.1	0.000	0.018	0.036	0.072	0.108	0.144	0.216
O2	Max	0.000	0.018	0.036	0.072	0.108	0.144	0.216	2.000
R1	Min	-0.1	0.000	0.004	0.008	0.015	0.023	0.030	0.045
HR	Max	0.000	0.004	0.008	0.015	0.023	0.030	0.045	2.000
R1	Min	-0.1	0.000	0.020	0.041	0.081	0.122	0.162	0.243
TOT	Max	0.000	0.020	0.041	0.081	0.122	0.162	0.243	2.000
R2	Min	-0.1	0.000	0.100	0.250	0.500	0.750	0.900	1.000
	Max	0	0.100	0.250	0.500	0.750	0.900	1.000	2.000
ACE	Min	-0.1	0	3	5	8	10	12	15
	Max	0	3	5	8	10	12	15	50
BPLV	Min	-1	60	110	150	170	200	-	-
	Max	60	110	150	170	200	400	-	-

After being accomplished the transformation tasks, the DM input table (DMIT) was created. Finally, in order to prepare the DM scenarios, the values were organized by hour considering the following variables:

- **Hour:** The hour associated to the values collected, all models use this variable;
- **Case Mix (CM)** – Age, admission type, admission provenance, Risk Patient, Insufficiencies Cardiac, CVA, Transplant, Surgical Admission;
- **SOFA** – Cardiovascular, Respiratory, Renal, Hepatic, Hematologic;

- **Accumulated Critical Events (ACE)** – ACE of Blood Pressure (BP), ACE of Oxygen Saturation (SPO2), ACE of Heart Rate (HR) and Total ACE;
- **Ratios 1 (R1)** – ACE of BP / max number of ACE of BP, ACE of SPO2 / max number of ACE of SPO2, ACE of HR / max number of ACE of HR;
- **Ratios 2 (R2)** – ACE of BP / elapsed time of stay, ACE of SPO2 / elapsed time of stay, ACE of HR / elapsed time of stay, Total of ACE / elapsed time of stay;
- **Ratios (R)** – Union of the two sets of ratios (R1 and R2).
- **Blood Pressure Last Values (BPLV)**: The last seven values collected in an hour, being one value for each column.

#### 4.4 Modelling

In this phase, the data mining models were induced using the data processed and transformed in the preceding phases. A set of Data Mining models (DMM) were induced using four DM techniques (DMT): GLM, SVM, DT and NB and two sampling methods: Holdout sampling (70% of the data for training and 30% for testing) and Cross Validation. Additionally the numeric attributes were used in two representation methods (Natural and Categorized). Scenarios 1 to 9 were manually configured. Scenario 10 (S10) was automatically configured by using variables selected by the DM engine (based in heuristic rules). A total of 160 models were induced.

$DMM = \{10 \text{ Scenarios}, 4 \text{ Techniques}, 2 \text{ Sampling Methods}, 2 \text{ Representation Methods}, 1 \text{ Target}\}$

##### Where the scenarios are:

S1: {CM, SOFA, ACE, R, RISK, BPLV}	S6: {CM, SOFA, ACE, R, RISK}
S2: {CM, SOFA, RISK, BPLV}	S7: {CM, RISK, BPLV}
S3: {CM, ACE, R, RISK, BPLV}	S8: {CM, SOFA, ACE, R, BPLV}
S4: {CM, RISK, BPLV}	S9: {CM, ACE, R, RISK, BPLV}
S5: {CM, R, RISK, BPLV}	S10: {Automatic}

##### Sampling Methods:

SM1: Holdout Sampling (HS)  
SM2: Cross Validation (CV)

##### Representation Methods:

RM1: Natural values (NAT)  
RM2: Categorized values (CATEG)

##### Techniques:

T1: Generalized Linear Models (GLM)	T3: Decision Trees (DT)
T2: Support Vector Machine (SVM)	T4: Naïve Byes (NB)

**Target:** TT1: Blood Pressure Critical Event

All the models were automatically induced in real-time, using streamed data and online-learning. A DM Model (DMM) can be represented by the following tuple:

$DMM = \langle \Delta, \alpha, DMT, DMSM, DMRM, DMTG, Hour, Age, AdmissionFrom, AdmissionType, Risk, ace\_bp, ace\_bp\_time, ace\_bp\_max, ace\_hr, ace\_hr\_time, ace\_hr\_max, ace\_spo2, ace\_spo2\_time, ace\_spo2\_max, total\_ace, total\_ace\_time, total\_ace\_max, SOFA\_Respiratory, SOFA\_Cardiovascular, SOFA\_Hepatic \rangle$



*SOFA\_Renal, SOFA\_Hematologic, BPLV1, BPLV2, BPLV3, BPLV4, BPLV5, BPLV6, BPV7*>

Where,

$\Delta$  is the DM rules and  $\alpha$  is the DM model configuration,  
*DMSM* is the sampling method and *DMRM* is the representation method,  
*DMT* is the DM technique and *DMTG* is the target,  
*Hour ... BPLV7* are the variables used by each model

For example if the Model 6 is composed by S6 using SVM, cross validation and categorized values, the tuple can be represented as:

*DMM6* = < $\Delta$ ,  $\alpha$ , *SVM, SM2, RM2, TTI, Hour, Age, AdmissionFrom, AdmissionType, Risk, ace\_bp, ace\_bp\_time, ace\_bp\_max, ace\_hr, ace\_hr\_time, ace\_hr\_max, ace\_spo2, ace\_spo2\_time, ace\_spo2\_max, total\_ace, total\_ace\_time, total\_ace\_max,, SOFA\_Respiratory, SOFA\_Cardiovascular, SOFA\_Hepatic, SOFA\_Renal, SOFA\_Hematologic*>

The configurations presented in table 5 were used to induce Data Mining models.

**Table 5.** Algorithms configurations

<i>Algorithm Name</i>	<i>Configuration</i>	<i>Value</i>
Generalized Linear Model	Automatic Preparation	On
	Confidence Level	0.95
	Enable Ridge Regression	Enable
	Missing Values Treatment	Mean for Numeric Mode for Categorical
	Ridge Parameter	10.0000
	Variance Inflation Factor	Disable
Support Vector Machine	Active Learning	Enable
	Automatic Preparation	On
	Complexity Factor	0.165605
	Kernel Function	Linear
	Tolerance	0.001
Decision Tree	Automatic Preparation	On
	Criteria For Splits	20
	Criteria For Splits (%)	0.1
	Maximum tree depth	7
	Minimum Child Record Count	10
	Minimum Records Per Node	0.05
	Tree Impurity Metric	Gini
Naive Bayes	Automatic Preparation	On
	Pairwise, Singleton Threshold	0

#### 4.5 Evaluation

After inducing all the models, the results achieved by each one of the 160 models were assessed. To assess the models the confusion matrix (CMX) was used. CMX allows determining the number of True Positives (TP) (predicted 1 and real 1), False Positives (FP) (1, 0), True Negatives (TN) (0, 0) and False Negatives (FN) (0, 1). Using the CMX is possible calculate some measures: Sensitivity =  $TP / (TP+FN)$ ; Specificity:  $TN / (TN+FP)$ ; Accuracy:  $TP / (TP+FP+TN+FN)$  and Precision:  $TP / (TP + FP)$

Table 6 presents the best results achieved by each technique and measure. For example, in the case of GLM the best accuracy (72.23%) and the best sensitivity was achieved by the same scenario (S3) but using different representation methods. The best sensitivity (95.90%) was achieved by the scenario 7 using SVM. The type of Sampling Methods used are not mentioned because the results are similar for both cases.

**Table 6.** Best model for each technique and measure

<i>Technique</i>	<i>Sensitivity</i>		<i>Specificity</i>		<i>Accuracy</i>	
GLM	S3RM2	0.9127	S2RM2	0.3866	S3RM1	0.7223
SVM	S7RM2	0.9590	S6RM2	0.2942	S5RM2	0.5586
DT	S6RM2	0.9316	S6RM2	0.3851	S6 S7 RM2	0.7152
NB	S3RM2	0.9100	S2RM2	0.3829	S7RM2	0.7099

To choose the best model, a threshold was introduced. The threshold combine three metrics in order to find the most suitable model to predict the probability of having a critical event (sensitivity) with an acceptable accuracy and precision in order to avoid a high number of false positives. The threshold defined was: Sensitivity  $\geq 90\%$  and Accuracy and Precision  $\geq 70\%$

Table 7 presents the best three models which achieved the threshold defined. The ranking was defined according to the sensitivity result.

**Table 7.** Best models achieving the threshold

<i>Model</i>	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>	<i>Precision</i>
S3TIRM2	0.7120	<b>0.9127</b>	0.3792	0.7310
S2TIRM2	0.7038	0.9125	0.3781	0.7346
S5TIRM2	0.7028	0.9120	0.3793	0.7359

## 5 Discussion

After analysing the models induced it was possible observe that the achieved results were not influenced by the sampling method used. Both approaches presented similar results. However the same consideration cannot be done concerning to the representation method. The achieved results by the models using categorized variables were significantly better. The models using natural values did not achieved the threshold. The model with a best sensitivity (95.90%) presented an accuracy of 49.26% using SVM and categorized variables. This model was rejected because did not achieved the threshold defined. In general, the models using categorized values increased the sensitivity in 10-15% and the accuracy in 1-3%.

Due the fact of the Intensivist preferring models sensitive to 1 (by predicting the worst scenario it is possible to avoid their occurrence. It is better predicting 1 and verify 0 than the opposite), the precision is a complementary measure to give an idea the number of correct predictions. It is important to note that to the intensivist is presented the probability (confidence rate) of a patient has a critical event in the next hours and not if he will have or not a critical event. It is not presenting a correct result (Yes or No) but the percentage of a critical event appear in the next hours (e.g. the probability of a patient has a BPCE in the next 10 hours is 80%). This reality decreases the

significance of false positives (FP), because the main goal is to avoid the occurrence of BPCE. Consequently the number of FP is higher.

In Figure 1 is possible observing the receiver operating characteristic (ROC) curve for the model which achieved the threshold and presented the best sensitivity (S3T1RM2). The curve is created by plotting the true positive rate (sensitivity) against the false positive rate (specificity) at various threshold settings.

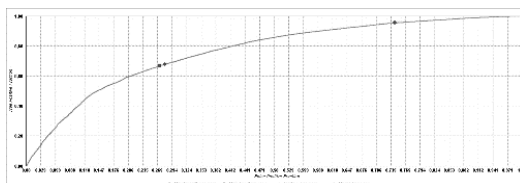


Figure 1. ROC for the most sensitivity model

## 6 Conclusions and Future Work

At the end of this work it was possible to assess the viability of using these variables to predict Blood Pressure Critical Events. The goal is not to predict effectively if the patient will have a BPCE or not but the probability of occurring a BPCE in the next hours. It was possible observe that the models using categorization methods presented best results. Only some of these models achieved the threshold defined. To the Intensivist the attained results (sensitivities between 90% and 95%) can help the decision and represents an important step in order to help to prevent possible cases of Hypertension or Hypotension. For them the number of False Positives (around 25%) are quite acceptable, once the goal is to prevent the occurrence of Critical Events. If the system is predicting a high probability of occurring a CE and the Intensivist can prevent it, the occurrence of a False Positive is an inevitable enjoyable. The models developed will be included in the INTCare ensemble engine. Scientifically the models produced can be used by other researchers in order to improve their works. The next step is focused in evaluating the predictions made by the system. In this process the intensivists will assess the clinical results. In addition this concept will be explored to other type of critical events.

**Acknowledgments.** This work has been supported by FCT - Fundação para a Ciência e Tecnologia within the Project Scope UID/CEC/00319/2013 and the contract PTDC/EEI-SII/1302/2012 (INTCare II).

## References

1. Kannel, W.B.: Risk stratification in hypertension: new insights from the Framingham Study. *American Journal of Hypertension* 13, 3S-10S (2000)
2. Portela, F., Santos, M.F., Machado, J., Abelha, A., Silva, Á., Rua, F.: Pervasive and intelligent decision support in Intensive Medicine—the complete picture. *Information Technology in Bio-and Medical Informatics*, pp. 87-102. Springer (2014)

3. Oliveira, S., Portela, F., Santos, M.F., Machado, J., Abelha, A., Silva, Á., Rua, F.: Predicting Plateau Pressure in Intensive Medicine for Ventilated Patients. *New Contributions in Information Systems and Technologies*, pp. 179-188. Springer (2015)
4. Filipe Portela, M.S., Jose Machado, Antonio Abelha, Álvaro Silva, Fernando Rua: Preventing Patient Cardiac Arrhythmias by Using Data Mining Techniques. In: 2014 IEEE Conference on Biomedical Engineering and Sciences. (2014)
5. Pedro Braga, Portela, F., Santos, M.F.: Data Mining Models to Predict Patient's Readmission in Intensive Care Units. In: ICAART - International Conference on Agents and Artificial Intelligence. (2015)
6. Veloso, R., Portela, F., Santos, M., Machado, J.M.F., Abelha, A., Silva, Á., Rua, F.: Real-time data mining models for predicting length of stay in intensive care units. (2014)
7. Portela, F., Santos, M.F., Machado, J., Abelha, A., Silva, Á.: Pervasive and Intelligent Decision Support in Critical Health Care Using Ensembles. *Information Technology in Bio-and Medical Informatics*, pp. 1-16. Springer Berlin Heidelberg (2013)
8. NIH, <http://www.nhlbi.nih.gov/health/health-topics/topics/hbp>, accessed in May, 2015
9. Wolf-Maier, K., Cooper, R.S., Banegas, J.R., Giampaoli, S., Hense, H.-W., Joffres, M., Kastarinen, M., Poulter, N., Primatesta, P., Rodríguez-Artalejo, F.: Hypertension prevalence and blood pressure levels in 6 European countries, Canada, and the United States. *Jama* 289, 2363-2369 (2003)
10. Mancia, G., Fagard, R., Narkiewicz, K., Redon, J., Zanchetti, A., Böhm, M., Christiaens, T., Cifkova, R., De Backer, G., Dominiczak, A.: 2013 ESH/ESC Guidelines for the management of arterial hypertension. *European Heart Journal* eht151 (2013)
11. Mahfoud, F., Ukena, C., Kandolf, R., Kindermann, M., Böhm, M., Kindermann, I.: Blood pressure and heart rate predict outcome in patients acutely admitted with suspected myocarditis without previous heart failure. *Journal of hypertension* 30, 1217-1224 (2012)
12. Mitsa, T.: *Temporal data mining*. CRC Press (2010)
13. Bucknall, T.K.: Medical error and decision making: learning from the past and present in intensive care. *Australian critical care* 23, 150-156 (2010)
14. Portela, F., Santos, M.F., Silva, Á., Machado, J., Abelha, A.: Enabling a Pervasive approach for Intelligent Decision Support in Critical Health Care. In: publication, S.-A.t. (ed.) *HCist 2011 – International Workshop on Health and Social Care Information Systems and Technologies*, pp. 10, Algarve, Portugal (2011)
15. João M. C. Gonçalves, Filipe Portela, Manuel F. Santos, Álvaro Silva, José Machado, Abelha, A.: Predict Sepsis Level in Intensive Medicine - Data Mining Approach. In: Springer (ed.) *Advances in Intelligent Systems and Computing*. Springer (2013)
16. Sérgio Oliveira, F.P., Manuel Filipe Santos, José Machado, António Abelha, Álvaro Silva and Fernando Rua: Intelligent Decision Support to predict patient Barotrauma risk in Intensive Care Units. In: Elsevier (ed.) *Procedia Computer Science - HCIST 2015 - Healthy and Secure People*, (2015)
17. Veloso, R., Portela, F., Santos, M.F., Silva, Á., Rua, F., Abelha, A., Machado, J.: A clustering approach for predicting readmissions in intensive medicine. *Procedia Technology* 16, 1307-1316 (2014)
18. Filipe Portela, R.V., Sérgio Oliveira, Manuel Filipe Santos, António Abelha, José Machado, Álvaro Silva and Fernando Rua: Predict hourly patient discharge probability in Intensive Care Units using Data Mining. *ScienceAsia Journal (ICCSCM 2014)* (2014)
19. Silva, A., Cortez, P., Santos, M.F., Gornesc, L., Neves, J.: Mortality assessment in intensive care units via adverse events using artificial neural networks. *Artificial Intelligence in Medicine* 36, 223-234 (2006)
20. Portela, F., Gago, P., Santos, M.F., Machado, J., Abelha, A., Silva, Á., Rua, F.: Pervasive real-time intelligent system for tracking critical events in intensive care patients. (2013)
21. Silva, Á., Cortez, P., Santos, M.F., Gomes, L., Neves, J.: Rating organ failure via adverse events using data mining in the intensive care unit. *Artificial Intelligence in Medicine* 43, 179-193 (2008)
22. Lee, P.: Design Research: What Is It and Why Do It? In: Reboot (ed.) *The Reboot*, vol. 2013, <http://thereboot.org> (2012)
23. Lunenfeld, P., Laurel, B.: *Design research: Methods and perspectives*. MIT press (2003)
24. Vincent, J.L., Moreno, R., Takala, J., Willatts, S., De Mendonca, A., Bruining, H., Reinhart, C.K., Suter, P.M., Thijs, L.G.: The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive care medicine* 22, 707-710 (1996)
25. Guy, W.: *ECDEU assessment manual for psychopharmacology*. Rockville, Md (1976)
26. W. Guy Modified From: Rush J, e.a.: Clinical Global Impressions (CGI) Scale. *Psychiatric Measures*, APA (2000)