Pervasive Ensemble Data Mining Models to Predict Organ Failure and Patient Outcome in Intensive Medicine

Filipe Portela[1], Manuel Filipe Santos[1], Álvaro Silva[2], António Abelha[3], and José Machado[3]

[1] Centro Algoritmi, University of Minho, Portugal
[2] Serviço de Cuidados Intensivos, Centro Hospitalar do Porto, Portugal
[3] CCTC, University of Minho, Portugal
{cfp,mfs}@dsi.uminho.pt,
moreirasilva@gmail.com,
{jmac,abelha}@di.uminho.pt

Abstract. The number of patients admitted to Intensive Care Units with organ failure is significant. This type of situation is very common in Intensive Medicine. Intensive medicine is a specific area of medicine whose purpose is to avoid organ failure and recover patients in weak conditions. This type of problems can culminate in the death of patient. In order to help the intensive medicine professionals at the exact moment of decision making, a Pervasive Intelligent Decision Support System called INTCare was developed. INTCare uses ensemble data mining to predict the probability of occurring an organ failure or patient death for the next hour. To assure the better results, a measure was implemented to assess the models quality. The transforming process and model induction are both performed automatically and in real-time. The ensemble uses online-learning to improve the models. This paper explores the ensemble approach to improve the decision process in intensive Medicine.

Keywords: Data Mining, Intensive Care, Organ Failure, Patient Outcome, INTCare, Ensemble, Real-time, Pervasive Health Care.

1 Introduction

This work is the culmination of the progress achieved in the INTCare project - an Intelligent Decision Support System (IDSS) for Intensive Medicine (IM). The first approach used offline-learning [1] and some patient data to predict organ failure and patient outcome with a good *accuracy* [1, 2]. To shift this concept to a real environment using real data and in real-time was a huge challenge.

Further work used some data acquired automatically, however the data was processed and transformed manually [3]. The results achieved were interesting; even though the agility of the process wasn't very good due to the high number of tasks which required human efforts. At this level, the solution required a changing on the

A. Fred et al. (Eds.): IC3K 2012, CCIS 415, pp. 410–425, 2013.

environment [4, 5] and on the way of the data is collected, processed and transformed. Using a set of intelligent agents [6] the entire Knowledge Discovery in Database (KDD) process [7] was automated. Nowadays the system prepares the data in real-time for the data mining tasks.

More recently, real-time data acquisition has been combined with online-learning [8]. The results were assessed in terms of *sensitivity, specificity* and *accuracy*. In the opinion of the physicians using only one measure wasn't the best option. Also, the way of how the models were compared is not the best. In order to overcome this problem, ensemble Data Mining (EDM) techniques were adopted and a data quality measure combining *sensitivity*, *accuracy* and *total error* was introduced. This paper presents the latest results. After a first experience where only 129 patients were evaluated a second one was made using more patient data, in this case the data comprises 335 patients. The main objective of this work is to understand if increasing the number of data available also increases the models' quality, in parallel, the benefits of using ensembles are studied.

This paper is divided into seven sections. After an introduction of the subject the principal concepts and the related work are described. The sections three and four present the initial phases of Knowledge Discovery in Database. Next, the fifth section presents the pervasive ensemble data mining and the sixth chapter makes an evaluation of the results achieved. Finally, some conclusions are outlined.

2 Background

2.1 Offline Vs. Online

The previous results obtained using EURICUS database [1] was the main motivation to develop this work. In EURICUS based work the variables used were collected manually in an offline mode: "The data was monitored, collected and registered manually, every hour, all Intensive Care Units (ICU) patient biometrics were recorded in a standardized sheet form by the nursing staff. The adverse events were also assigned in a specific sheet at an hourly basis."[1]. The variables used were: Age, Critical Events, Admission Variables, Outcome, and Sepsis-related Organ Failure Assessment (SOFA) [9, 10].

Now, the objective is to obtain all of those variables automatically and at same time to induce data mining models using an online approach in order to predict the organ failure and patient outcome in real-time. The greater challenge is the development of some procedures using all the values obtained by the data acquisition system instead of using hourly values. This change allows for a continuous data monitoring.

2.2 INTCare System

INTCare system is composed by four subsystems [11]: Data acquisition, knowledge management, Inference and Interface. Each subsystem is autonomous and uses intelligent agents [12] to perform automatically some tasks. For data acquisition were

used the following agents: Gateway, Vital Signs Acquisition agent, Electronic Nursing Record (ENR) agent, Laboratory Results (LR) agent and Electronic Health Record (EHR / AIDA). Furthermore, it was used a pre-processing agent for the data validation and data transformation. Finally, the induction of data mining models is ensured by the Data Mining (DM) agent. The ensemble is induced automatically and in real-time by DM agent, whenever some request is done.

2.3 Knowledge Discovery Process

Knowledge Discovery from Databases (KDD) process is recognized as a process which can obtain new knowledge using some data. This process is composed by five stages: Selection, Pre-Processing, Transformation, Data Mining and Interpretation [7]. Figure 1 shows the KDD process for the ICU data. The database is populated with data from seven major sources. The data are selected from the data warehouse to be processed or transformed, depending on the goal of each one of the variables. After this task, the data are available to be presented by the Electronic Nursing Record (ENR) and prepared for creating Data Mining Models. Finally, all models are evaluated and the obtained knowledge is presented in the INTCare system.

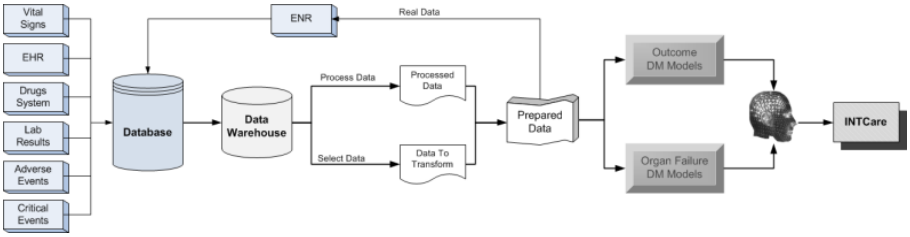

Fig. 1. ICU Knowledge Discovery in Database Process

2.4 Ensemble Data Mining

The use of Data Mining (DM) techniques in the medical area has been gaining an increasing interest by researchers [13]. Being this work a typical DM classification problem [14] and bearing in mind the idea of having a pervasive and real-time IDSS, a set of DM approaches were explored.

In the last experience ensemble data mining has been adopted. The reason of using ensembles has to do with the established principle that: the *sensitivity* can usually be improved by using ensembles of predictive models instead of a single model [15]. The ensemble-learning methodology consists in two sequential phases: the training and the testing phase [16]. In the training phase several different predictive models are generated from the training set. In the test phase the ensemble is executed and aggregates the outputs for each predictive model [16]. In this project was followed the Stacked Generalization methodologies [17, 18] and the learning procedure was divided into four steps [16]. To split the dataset the stratification technique is used. For each target a different dataset was considered with the same distribution on the target classes (0, 1).

The use of Oracle technology, especially Oracle Data Mining (ODM) facilitated the ensemble induction [19, 20]. Classification techniques used are some of the most used in DM [21] such as Support Vector Machine (SVM), Decision Trees (DT) and Naïve Byes (NB).

2.5 Pervasive Health Care

During the development of INTCare, some features were added according to the pervasive health care concept [22, 23], having as main purpose making the system available anywhere and anytime [5]. According to Varshney [24] pervasive health care can be defined as "conceptual system of providing healthcare to anyone, at any time, and anywhere by removing restraints of time and location while increasing both the coverage and the quality of healthcare". This approach is based on information that is stored and available online [25]. In order to turn the system into pervasive system pervasive computing features should be considered. Satyanarayanan [26] characterizes the pervasive computing as an evolutionary step resulting from previous two steps: first distributed computing and then mobile computing. The main characteristics are smart spaces, invisibility, localized scalability and uneven conditioning. INTCare uses pervasive computing features and it is fed by the probabilities provided by the ensemble data mining process. The results attained are displayed in situated devices.

2.6 Related Work - Forth Approach (Ensemble Data Mining)

The models were induced in real-time using online-learning by the DM agent [27]. To evaluate the ensemble three measures were considered: *Sensitivity, Accuracy,* and *Total Error (Terror).* The average and the standard deviation of each one of the measures were estimated considering 10 runs. The use of ensemble helps to select the best model in the cases where more than one model present good results (e.g. outcome, hepatic and the respiratory systems). For each target / fold a separate dataset has been considered. The data corresponds to:

- Period in analysis: 105 days;
- Number of patients: 129.

Table 1. Ensemble Results

Target	*Accepted by quality measures*	*Sensitivity*	*Accuracy*	*Specificity*	*Terror*
Cardiovascular	YES	**97,95** ± 0,31	**76,81** ± 2,35	**41,81** ± 5,75	**23,19** ± 2,35
Coagulation	YES	**91,20** ± 3,57	**65,69** ± 3,83	**49,61** ± 6,15	**34,31** ± 3,84
Hepatic	NO	**69,24** ± 9,41	**82,89** ± 2,57	**87,34** ± 3,22	**17,10** ± 2,57
Outcome	YES	**99,77** ± 0,33	**63,58** ± 3,11	49,58 ± 4,90	**36,42** ± 3,11
Renal	NO	**77,17** ± 12,41	**43,08** ± 4,66	**43,08** ± 4,66	**49,09** ± 5,39
Respiratory	NO	**67,11** ± 5,67	**63,86** ± 4,27	**60,39** ± 6,75	**36,14** ± 4,27

Table 1 presents the performance achieved by the ensemble for each target. The values correspond to the average of the measures obtained during ten runs of the ensemble. Each average has associated the standard deviation. Respiratory, hepatic and renal systems don't meet the measures established and aren't considered by the pervasive system.

3 Data Selection and Pre-processing

The two initial phases of Knowledge Discovery in Database (KDD) process [4] use the data acquisition system presented in the background section (Figure 1) to obtain the data. The first phase is concerned to the data selection from database and it is in agreement with data necessary to feed the DM Models:

ICU_HL7 ⊆{Vital Signs}
ICU_HL7_T ⊆{Vital Signs auto validated (real values)}
ICU_PARAM ⊆{ICU Limits (max, min) values}
ICU_LR ⊆{All Lab Results}
ICU_DRUGS ⊆{All Patient Drugs administrated}
ICU_ENR ⊆{Data validated and provided from ENR}
ICU_CEVENTS ⊆{ICU Critical Patient Events}
EHR_ADMIN ⊆{ICU Patient Admission}
EHR_OUT ⊆{ICU Patient Outcome}

The second phase is responsible for the automatic data validation and patient identification. In this phase it is ensured that all data collected are valid and are correctly identified, i.e. all values collected are within the normal ranges of ICU values, and they have a valid patient identification (PID) [6].

At pre-processing phase, other procedures are executed to prepare the Data Mining input table. For instance, only the values collected during the first five days are used. When the patient is admitted into the ICU, an agent prepares automatically the table adding 120 rows for that patient. When the patient goes out, if he/she leaves before 120 hours, the rows in excess are deleted. This table is used as a temporary table for DM input. This table stores: the case mix values for all 120 lines (hours), the number of Critical Events and SOFA (0, 1) values for each hour. DM agent gets all variables present in temporary tables and, in addition, calculates the values in fault.

4 Transformation

The third phase of the KDD process is autonomous requiring no manual actions. All the tasks are performed automatically and in real-time by the INTCare intelligent agents. The variables in use are:

SOFA $_{\text{Cardio, Respiratory, Renal, Liver, Coagulation, neurologic}}$ = {0, 1}

Case Mix = {Age (1-4), Admission type (U or P), Admission from (1-6)}

Critical Events Accumulated (ACE) = {ACE of Blood Pressure (BP) (lN), ACE of Oxygen Saturation (SpO2) (lN), ACE of Heart Rate (HR) (lN), ACE of Urine Output (Ur) (lN)}

Ratios1 (R1) = {ACE of BP/elapsed time of stay (Q+), ACE of SO2/elapsed time of stay (Q+), ACE of HR/elapsed time of stay (Q+), ACE of Ur/elapsed time of stay (Q+), Total of ACE / elapsed time of stay (Q+)}}

Ratios2 (R2) = {ACE of BP / max number of ACE of BP (Q+), ACE of SO2/ max number of ACE of SO2 (Q+), ACE of HR / max number of ACE of HR (Q+) , ACE of Ur / max number of ACE of Ur (Q+), Total of ACE (Q+), Total of ACE / Total ACE max (Q+)}

Ratios (R) = R1 U R2

ACE of HR. Sum of values in a hourly base for each event type, i.e. if in the first hour has 1 event and in the second hour 2, the ACE for the second hour is 3.

Total of ACE is the sum of all ACE for the hour.

Max Number of ACE is the maximum number of each variable present in Table 5.

Elapsed Time of Stay. Total number of hours elapsed since the patient admission in the moment when ratio is calculated.

Total ACE max is the maximum value for ACE obtained by a patient in a specific hour (Table 5).

Outcome = {0, 1}

Table 2 presents the values considered by Data Mining models.

As referred in the previous work [8], the first transformation process is a simple task for analysing the values collected and for transforming them according to some rules (if then else). This process is applied to the variables presented in Table 2. When there is a case mix, all variables are inserted in the database. When a patient comes into the ICU, a procedure is executed. Regarding to the age parameter, the procedure verifies the patient age. For the admission type and origin, the admission form is consulted in the EHR. In all the cases the values are processed and the value is inserted into DM_INPUT table. In the case of the SOFA, the approach is a little bit different, the values are collected in real-time and in a continuous way.

The data mining models only use one value per hour. All collected values are considered and then the final value is assigned. If it is verified more than one result by hour, only the worst value of the hour it is considered. For example, in the case of cardiovascular system, there are five different possibilities to be categorized as 1 (BP, Dopamine, Dobutamine, Epi and Norepi). The SOFA values are then transformed into binary variables, where 0 describes normality and 1 describes dysfunction/failure and comprises the original SOFA value. By default the SOFA value variable is 0 and, if some condition is verified (e.g. for coagulation, platelets <=150) the values are

Table 2. Variables transforming (example)

ID		*Variable*	*Min*	*Max*	*Value*
Age		-	18	46	1
		-	47	65	2
		-	66	75	3
		-	76	130	4
Admission Type		Urgent	-	-	u
		Programed	-	-	p
Admission From		Chirurgic	-	-	1
		Observation	-	-	2
		Emergency	-	-	3
		Other ICU	-	-	4
		Other Hospital	-	-	5
		Other Situation	-	-	6
SOFA	Cardio	BP (mean)	0	70	1
		Dopamine	0,01	-	1
		Dobutamine	0,01	-	1
		Epi / Norepi	0,01	-	1
	Renal	Creatinine	1.2	-	1
	Resp	Po2/Fio2	0	400	1
	Hepatic	Bilirubin	1.2	-	1
	Coagul	Platelets	0	150	1
	Neuro	Glasgow	3	14	1

updated to 1. This update has effect in the starting date when the value was measured. The outcome value (live or died) is updated according to the patient discharge condition, when the patient die all of the values in the input table are updated to 1.

The second transformation phase uses critical events. Firstly, a set of procedures are executed in order to understand if a value is critical and if the event is adverse. Table 3 presents the variables in study and the *min* and *max* values for each case.

Table 3. Data Ranges

EvId	*Descr*	*Min EC*	*Max EC*	*Min Val*	*Max Val*	*Min Any*	*Max Any*
1011	BP	90	180	0	300	60	
3000	O2	90	100	0	100	80	
2009	HR	60	120	0	300	30	180
DIU	UR	30	1000	0	1000	10	

The values are validated and then the system determines if they are critical and how critical are they. According to Table 4, a value can be considered *normal* (0), *critical* (1) or *too critical* (2). If a value is critical (1), the event will be considered critical only if the values collected are maintained during a period of time. If the value is spontaneous and is too bad (2), the event will be always considered critical independently the extent of the event. Then, each collected value will be inserted in the events table according to the event type and if the predecessor event is or not the same. To know if this event type is of the same type a flag is used. A procedure is used to understand if the critical values collected may or may not represent a critical event. To this end the Table 4 is used. To consider an event as a critical event two

Table 4. The protocol for the out of range physiologic measurements (adapted by Álvaro [1])

	BP (mmHg)	*SpO2 (%)*	*HR (bpm)*	*UR (ml/h)*
Normal range	90 - 180	>= 90	60 - 120	>= 30
Critical event [a]	>= 1h	>= 1h	>= 1h	>= 2h
Critical event [b]	< 60	<80	<30 V> 180	<= 10

[a] Defined when continuously out of range.

[b] Defined anytime.

main constraints should be satisfied. For example, in the case of SpO2 the values should be between 80 and 90 for more than one hour, or less than 80 during some period.

The next procedure reviews the values according to the event time and event type: if it is verified a critical event, the value will be inserted into the critical events table. Then a procedure is executed hourly. This procedure calculates the Accumulated Critical Events (ACE) – to reflect the patients' clinical evolution/severity of illness by hour.

The next step consists in obtaining the ACE ratios. To implement this process a set of calculations are executed in the exact moment when the value was collected. This process requires more memory and processing time and can delay the other procedures. For the ratios is used the maximum number of occurrences verified in the past to a specific hour (Table 5). The maximum values are updated according to the maximum number of events verified for a patient for each variable in a specific hour.

The maximum number of *ACE* is automatically determined by hour. A procedure is executed to verify if the number of *ACE* is higher than the values present in the database. Table 5 presents an example of the maximum *ACE* values for each variable / hour verified in a patient. For example, in the case of blood pressure the maximum number of *ACE* observed in a patient at the 20th hour is six. However, at the 30th hour it is five. This result signifies that the patient who had a bigger number of BP CE at the 20th hour was discharged before the 30th hour of hospitalization. The total events are the maximum number of events occurred for all the categories and verified only in a patient during a specific hour.

Table 5. Critical Events daily number (example)

Variables	***Max number of accumulated critical events***										
Hour	***1***	***2***	***3***	***4***	***5***	***10***	***15***	***20***	***30***	***40***	***50***
blood pressure events	0	1	1	1	1	1	3	6	5	3	3
heart rate events	0	1	1	1	2	2	4	3	3	7	8
oxygen events	0	2	2	2	2	4	6	7	9	10	11
urine events	0	0	0	0	0	0	0	0	0	0	0
Total events	0	2	2	2	3	6	8	10	16	15	16

The next procedure calculates the ACE and all ratios for DM model. During all the processes described above, a procedure is responsible to get all data generated and store them into a specific table for the DM task. Finally, and after having all values correctly inserted into DM input table, another procedure cleans the inconsistent values. This procedure is responsible to delete all rows having null / incorrect values.

Table 6 presents the discretization rules defined to code continuous values (values ϵ {ℝ0+}). The ranges were created using a 7-point-scale adapted from Clinical Global Impression - Severity scale (CGI-S) [28]. The goal of *CGI-S* is to allow the clinician to rate the severity of illness [29]. The boundaries of each set were defined by the ICU doctors considering the significance of each set of values. More severe cases are assigned to levels 6 and 7. At the top of the table is the identification of the set. The left column identifies the variable. In the middle of the table are defined the ranges for each set.

The *R1min* and *R1max* are used by *R1* (max number of ACE). According to the percentage of the value it is categorized. For example, for the *R1* attribute '*ACE of BP / max number of ACE of BP*', if a patient has 7 *ACE* at the sixth hour and the maximum verified in the past for this time is 10 *ACE*, the respective set, according to table IV, will be 4 (7/10 = 0,7).

For all attributes of R1, the ranges of the set are equal. For the *R2* (ratios that use the elapse time) attributes, it is used the rows (*R2 BP min to R2 TOT max*) to determine the set. In this case, each attribute has a different range. For the attribute '*ACE of O2/elapsed time of stay'* if, for example, a patient has 10 ACE of O2 at the 50th hour, the ratio value will be 0,2 and the DM set will be 5 (0,1<0.2<=0,3).

Finally, all ACE values are grouped in accordance with their importance and number. For example, a patient that has 8 *ACE*, using table 6 and ace row, corresponds to the set 3 (5 < 8 <=8) in the DM input table.

Table 6. Discretization set of Data Mining Inputs

SET		0	1	2	3	4	5	6	7
R1	Min	-0,1	0	0,2	0,4	0,6	0,8	1	-
	Max	0	0,2	0,4	0,6	0,8	1	+00	-
R2	Min	-0,1	0,000	0,020	0,040	0,075	0,100	0,300	0,500
BP	Max	0,000	0,020	0,040	0,075	0,100	0,300	0,500	1
R2	Min	-0,1	0,000	0,020	0,040	0,075	0,100	0,300	0,500
O2	Max	0,000	0,020	0,040	0,075	0,100	0,300	0,500	1
R2	Min	-0,1	0,000	0,001	0,003	0,006	0,010	0,030	0,100
HR	Max	0,000	0,001	0,003	0,006	0,010	0,030	0,100	1
R2	Min	-0,1	0,000	0,020	0,050	0,080	0,100	0,300	0,500
UR	Max	0,000	0,020	0,050	0,080	0,100	0,300	0,500	1
R2	Min	-0,1	0,000	0,020	0,050	0,080	0,100	0,300	0,300
TOT	Max	0,000	0,020	0,050	0,080	0,100	0,300	0,500	1
ACE	Min	-0,1	0	3	5	8	10	12	15
	Max	0	3	5	8	10	12	15	50

5 Pervasive Ensemble Data Mining

Fig. 2 provides an overview of Data Mining modulation. In this figure it can be observed that the data preparation module is executed using the data stored in the database. After the transformation phase, data are stored into DM_INPUT_DB table. This table is then used to predict the value of the six target variables. Afterwards, the obtained results are applied into the prediction table of the patients admitted into ICU (UCI_PATIENT_5DAYS_AG). Then, six new columns are added containing the prediction of value 1 (occur an organ failure or patient die) for each target. In the data mining engine, each target constitutes an individual process.

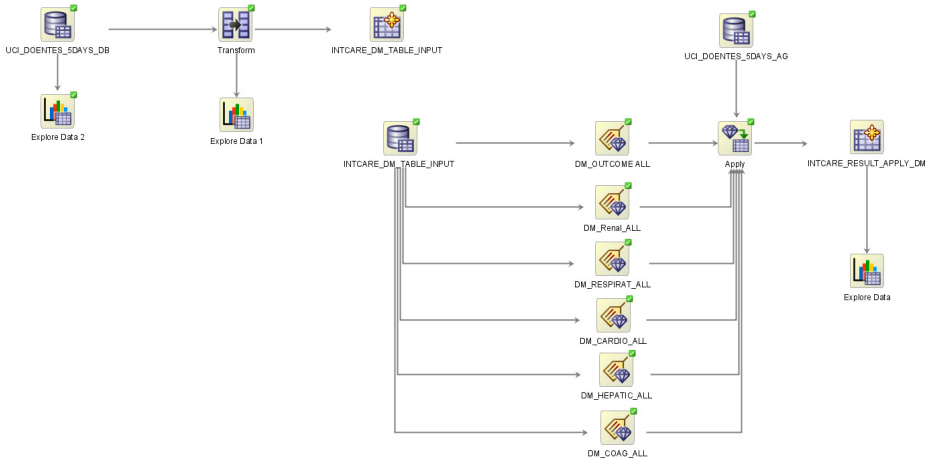

Fig. 2. DM Model

The fourth approach encloses the ensemble Data Mining techniques; in this case this model is executed 10 times for each target.

In order to automate this process, some researches has been done to find how to induce DM models automatically. As result it was possible to develop a procedure to execute the DM engine in real-time. The DM agent is responsible to run the engine whenever a request is made. The process of inducing Data Mining models is divided into two steps:

- Predictive Models – 126 models are induced combining seven scenarios (S1 to S7), six targets and three different techniques (SVM, DT and NB);
- Ensemble – the models are assessed in terms of the *sensitivity*, *accuracy, total error* and *specificity*. The best model for each target (t) is then selected.

The main purpose of the ensemble is to select the most suited model from a set of candidates. In order to evaluate the models, a quality measure was defined. This measure is based in the results obtained by the models in terms of *sensitivity*, *accuracy* and *total error*. The selected models are used by the pervasive system only if they satisfy the following conditions:

- *Total Error* <= 40%
- *Sensitivity* >= 85%
- *Accuracy* >= 60%

These thresholds were defined in order to assure a minimum level of quality in models. The measure was defined in accordance with ICU doctors. The values can be adjustable anytime and are commonly accepted in the medical community.

The ensemble can be defined as a three-dimensional matrix *M* composed by *s=7* scenarios (*s1* to *s7*) x *t=6* targets (*t1* to *t7*) x z=3 techniques (*z1* to *z3*). Each element of *M* corresponds to a particular model and can be defined as:

$$M_{s,t,z} = \begin{Bmatrix} s = 1 \dots 7 \\ t = 1 \dots 6 \\ z = 1 \dots 3 \end{Bmatrix}$$

Where,

s:	t:	z:
1 = {CASE MIX}	1 = Respiratory	1 = Support Vector Machine
2 = {CASE MIX, ACE, R}	2 = Cardiovascular	2 = Decision Trees
3 = {CASE MIX, ACE, R1}	3 = Coagulation	3 = Naïve Bayes
4 = {CASE MIX, ACE, SOFA}	4 = Renal	
5 = {CASE MIX, ACE, SOFA, R}	5 = Hepatic	
6 = {CASE MIX, ACE, SOFA, R2}	6 = Outcome	
7 = {CASE MIX, ACE, SOFA, R1}		

Each model is induced automatically and in real-time using streamed data. After all data have been processed, the models are induced. Here occurs the unique manual operation, i.e., the data mining models must be manually configured. In this phase 108 models were developed (*6 targets x 6 models x 3 techniques*). The data mining models were induced into two steps: the first is responsible to prepare the final data to be used by the prediction models; in the second, the data obtained by the first stage are used by DM techniques to predict the probability of failure of each organ and patient outcome.

In the first step, the data stored in the DM input table is loaded. The numbers (ACE and ratios) are distributed using the presented discretization techniques. The other values are maintained as they are, and a final table is generated. During the DM modulation the neurologic system was not considered due to the high number of data in fault.

6 Evaluation

In order to evaluate the models, a test phase using online-learning was performed. The Data Mining techniques were applied on the following dataset:

Data Description:

Collection Time:	367 days
Patients Number:	335
Data Considered:	Values of five first days
Exclusion criterion I:	Patient with data collected intermittently, i.e., the collection system failed at least more than one hour in a continuous way;
Exclusion criterion II:	Existence of null values;

6.1 Data Input Distribution of the Inputs and Targets

Tables 7 and 8 present the distribution of the attributes considered. For the test phase the original dataset (DMIT) was divided into two different datasets using the holdout sampling method: 70% of the data were considered for training and 30% for testing (stratified by the target). Each target has a different dataset.

Table 7. Input Case Mix Variables distribution

Attribute	1	2	3	4	5	6	U	P
AGE	12,40	41,10	18,85	27,65	-	-	-	-
ADMIN_FROM	50,69	0,22	16,81	11,86	2,21	18,21	-	-
ADMIN_TYPE	-	-	-	-	-	-	72,58	27,42

Table 8. Input ACE and Ratios Variables distribution (%)

Attribute	0	1	2	3	4	5	6	7
ACE_BP	82,06	14,56	1,75	0,91	0,31	0,06	0,24	0,11
ACE_BP_MAX	73,57	1,61	7,05	5,78	3,51	1,21	7,26	-
ACE_BP_TIME	82,06	1,33	4,74	5,79	2,49	1,30	1,24	1,05
ACE_HR	80,62	13,76	3,35	1,25	0,38	0,37	0,23	0,04
ACE_HR_MAX	80,62	1,78	5,98	6,75	2,12	0,74	2,01	-
ACE_HR_TIME	80,62	2,64	2,86	2,77	1,79	1,66	2,85	4,81
ACE_O2	73,44	16,51	3,63	2,77	1,13	1,01	0,38	1,12
ACE_O2_MAX	73,40	6,17	6,21	7,37	2,79	1,13	2,94	-
ACE_O2_TIME	73,39	3,76	5,42	6,00	3,92	2,51	2,53	2,46
ACE_UR	82,06	14,56	1,75	0,91	0,31	0,06	0,24	0,11
ACE_UR_MAX	82,06	1,33	4,74	5,79	2,49	1,30	1,24	1,05
ACE_UR_TIME	73,57	1,61	7,05	5,78	3,51	1,21	7,26	-
ACE_TOT	61,50	21,06	5,07	4,18	1,94	1,55	1,37	3,32
ACE_TOT_MAX	61,42	11,91	11,78	8,00	3,61	1,10	2,18	-
ACE_TOT_TIME	61,42	5,22	7,50	9,11	5,69	2,95	3,93	4,19

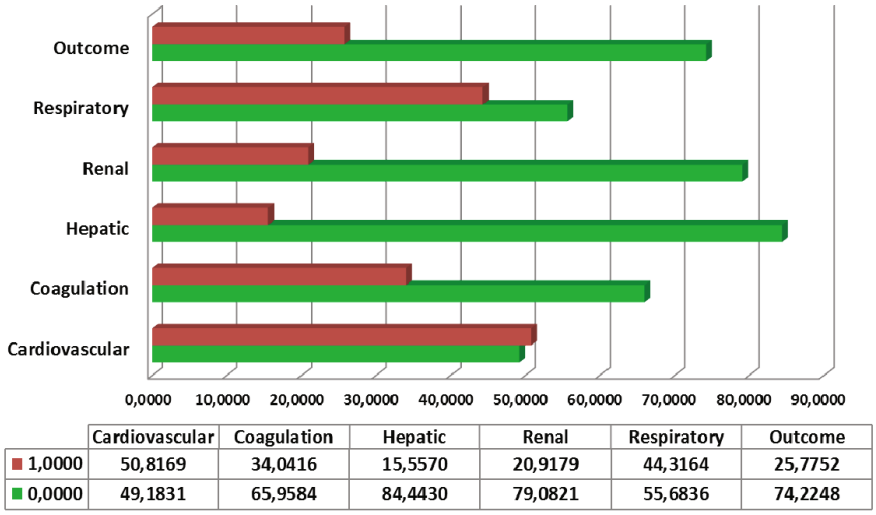

Fig. 3. Distribution of the classes

Fig. 3 presents the distribution of the classes (in percentage) for each one of the targets. For example, in the input dataset 15,56 % of the rows has the value of renal variable equal to 1. This value does not means that 25,78% of the patients has an renal failure, because the values are related to the rows and not with the patients.

This means that 25,78% of the records correspond to patients with renal failure. In this case, only the cardiovascular system present a high level of cases with organ failure (cardiovascular result = 1): 50,82%. Respiratory system presents a difference of 10% between the two values, in this case 44,32% of the records present a value equal to 1. In all of the other targets the numbers of positive cases (final result = 1) are substantially lower.

6.2 Ensemble Results

After the DM engine has been run, the best results obtained for each target are compared. Table 9 presents the performance achieved by the ensemble for each target. The values correspond to the average of the measures obtained during ten runs of the ensemble. The hepatic, respiratory and renal system didn't meet the measures established. In some executions of the renal system the results, reached the measure established, however having into account the average of the 10 runs, the results weren't satisfactory. Table 9 shows these results, as can be seen, the maximum *sensitivity* of renal system was 91,89.

Table 9. Ensemble Results by organ systems and outcome (%)

Target	*Sensitivity (avg)*	*Specificity (avg)*	*Accuracy (avg)*	*Terror (avg)*
Cardiovascular	**98,02**	35,52	**74,59**	**25,41**
Coagulation	**90,14**	46,47	**63,79**	**36,21**
Hepatic	57,34	84,18	**79,03**	**20,97**
Renal	71,50	42,44	49,21	50,79
Respiratory	67,03	42,09	59,26	40,74
Outcome	**97,30**	45,03	**62,35**	**37,65**
Target	*Sensitivity (max)*	*Specificity (max)*	*Accuracy (max)*	*Terror (max)*
Cardiovascular	99,60	43,72	77,80	27,79
Coagulation	94,62	60,56	73,22	41,59
Hepatic	79,69	93,38	83,77	25,12
Renal	91,89	54,70	57,18	60,48
Respiratory	73,89	67,50	69,50	52,56
Outcome	100,00	57,61	70,52	43,64

6.3 Comparing the Results

Comparing the results (table 10) previously obtained (table 1) with those observed in this study (table 9) it is possible to conclude that the difference is not significant. In general the results are widely satisfactory, for example the cardiovascular system presents a average *sensitivity* 0,07 upper, however the *accuracy* and *total error* are worst. The models are chosen according to the results obtained in terms of the measure, at the moment when data mining engine requires a model to answer to some target. Table 10 presents the variation (Δ) between the previous results and the newer.

Table 10. Variation of the obtained results

Target	*Sensitivity* OLD	*Accuracy* OLD	*TERROR* OLD	*Sensitivity* Δ	*Accuracy* Δ	*TERROR* Δ
Cardiovascular	97,95	76,81	23,19	0,07	-2,22	2,22
Coagulation	91,20	65,69	34,31	-1,06	-1,9	1,9
Hepatic	69,24	82,89	17,10	-11,9	-3,86	3,87
Renal	77,17	43,08	36,42	-5,67	6,13	14,37
Respiratory	67,11	63,86	49,09	-0,08	-4,6	-8,35
Outcome	99,77	63,58	36,14	-2,47	-1,23	1,51

For each one of the six targets was used three different DM techniques (SVM, DT and NB). Table 11 describes the experimental settings considered for each technique. In addition, this table indicates whether the used value corresponds to a default value or to a user-defined value.

Table 11. Techniques Configurations

DM Technique	Parameter Name	Parameter Value	Parameter Type
DT	Minrec Node	10	Input
	Max Depth	7	Input
	Minpct Split	.1	Input
	Impurity Metric	Gini	Input
	Minrec Split	20	Input
	Minpct Node	.05	Input
	Prep Auto	On	Input
NB	Pairwise Threshold	0	Input
	Singleton Threshold	0	Input
SVM	Conv tolerance	.001	Input
	Active learning	Al Enable	Input
	Kernel function	Linear	Default
	Complexity factor	0.142831	Default
	Prep auto	On	Input

7 Conclusions

The development of autonomous and real-time IDSS is a big challenge. The most difficult is the data transformation task (the process of obtaining critical events and their ratios). Ensemble data mining using online-learning and real-time data is another obstacle. This work goes in this direction, proving that it is possible to implement an IDSS in critical health environments minimizing the human intervention.

After automate the entire KDD process, in particular, the data transformation and the data mining processes, a new approach was experienced using ensembles.

INTCare is now an autonomous system and it is able to automatically and in real-time to predict the organ failure and patient outcome for the next 24 hours. The DM engine operates autonomously running the models and broadcasting the results through the INTCare system.

Experimental work was conducted in order to determine whether the use of online-learning and ensembles are or not a good option. The results are in favour of the last approach. The tasks can now be performed automatically and in real-time with a reduced human efforts. Additionally, the better models are used in the precise moment the ICU professionals need to take a decision. Furthermore, the performance of the models doesn't decrease over the time. The results observed in both situations were satisfactory corroborating the hypothesis of adopting ensemble approach.

Future work will include more data to optimize the DM models, in particularly the urine output and Glasgow Score.

Acknowledgements. This work is supported by FEDER through Operational Program for Competitiveness Factors – COMPETE and by national funds though FCT – Fundação para a Ciência e Tecnologia in the scope of the project: FCOMP-01-0124-FEDER-022674. The authors would like to thank FCT (Foundation of Science and Technology, Portugal) for the financial support through the contract PTDC/EIA/72819/ 2006 (INTCare) and PTDC/EEI-SII/1302/2012 (INTCare II). The work of Filipe Portela was supported by the grant SFRH/BD/70156/2010 from FCT.

References

1. 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)
2. 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)
3. Boas, M.V., Santos, M.F., Portela, F., Silva, A., Rua, F.: Hourly prediction of organ failure and outcome in real time in Intensive Care Medicine (2010)
4. Portela, F., Santos, M.F., Vilas-Boas, M.: A Pervasive Approach to a Real-Time Intelligent Decision Support System in Intensive Medicine. In: Fred, A., Dietz, J.L.G., Liu, K., Filipe, J. (eds.) IC3K 2010. CCIS, vol. 272, pp. 368–381. Springer, Heidelberg (2013)
5. Portela, F., Santos, M.F., Silva, Á., Machado, J., Abelha, A.: Enabling a Pervasive approach for Intelligent Decision Support in Critical Health Care. In: HCist 2011 – International Workshop on Health and Social Care Information Systems and Technologies, Algarve, Portugal, p. 10. S.-A.T. Publication (2011)
6. Portela, F., Santos, M.F., Gago, P., Silva, Á., Rua, F., Abelha, A., Machado, J., Neves, J.: Enabling real-time intelligent decision support in intensive care. In: 25th European Simulation and Modelling Conference- ESM 2011, 446 pages (2011)
7. Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P.: From data mining to knowledge discovery: an overview (1996)
8. Portela, F., Pinto, F.: Data Mining Predictive Models For Pervasive Intelligent Decision Support In Intensive Care Medicine. In: KMIS 2012 - nternational Conference on Knowledge Management and Information Sharing, Barcelona, INSTICC (2012)
9. Vincent, J., Mendonça, A., Cantraine, F., Moreno, R., Takala, J., Suter, P., Sprung, C., Colardyn, F., Blecher, S.: Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: Results of a multicenter, prospective study. Critical Care Medicine 26, 1793–1800 (1998)

10. 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)
11. Portela, F., Santos, M., Vilas-Boas, M., Rua, F., Silva, Á., Neves, J.: Real-time Intelligent decision support in intensive medicine. In: KMIS 2010- International Conference on Knowledge Management and Information Sharing, Valência, Espanha, p. 7 (2010)
12. Santos, M.F., Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., Neves, J.: INTCARE - Multi-agent approach for real-time Intelligent Decision Support in Intensive Medicine. In: 3rd International Conference on Agents and Artificial Intelligence (ICAART). Springer (2011)
13. Bellazzi, R., Zupan, B.: Predictive data mining in clinical medicine: current issues and guidelines. International Journal of Medical Informatics 77, 81–97 (2008)
14. Han, J., Kamber, M.: Data mining: concepts and techniques. Morgan Kaufmann (2006)
15. Dietterich, T.: Ensemble methods in machine learning. Multiple Classifier Systems, 1–15 (2000)
16. Kantardzic, M.: Data mining: concepts, models, methods, and algorithms. Wiley-IEEE Press (2011)
17. Ting, K.M., Witten, I.H.: Issues in stacked generalization. Arxiv preprint arXiv:1105.5466 (2011)
18. Wolpert, D.H.: Stacked generalization*. Neural networks 5, 241–259 (1992)
19. Concepts, O.D.M.: 11g Release 1 (11.1). Oracle Corp. 2007 (2005)
20. Tamayo, P., Berger, C., Campos, M., Yarmus, J., Milenova, B., Mozes, A., Taft, M., Hornick, M., Krishnan, R., Thomas, S.: Oracle Data Mining. In: Data Mining and Knowledge Discovery Handbook, pp. 1315–1329 (2005)
21. Wu, X., Kumar, V., Ross Quinlan, J., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G.J., Ng, A., Liu, B., Yu, P.S.: Top 10 algorithms in data mining. Knowledge and Information Systems 14, 1–37 (2008)
22. Varshney, U.: Pervasive healthcare and wireless health monitoring. Mobile Networks and Applications 12, 113–127 (2007)
23. Orwat, C., Graefe, A., Faulwasser, T.: Towards pervasive computing in health care - A literature review. BMC Medical Informatics and Decision Making 8, 26 (2008)
24. Varshney, U.: Pervasive Healthcare Computing: EMR/EHR, Wireless and Health Monitoring. Springer-Verlag New York Inc. (2009)
25. Mikkonen, M., Väyrynen, S., Ikonen, V., Heikkila, M.O.: User and concept studies as tools in developing mobile communication services for the elderly. Personal and Ubiquitous Computing 6, 113–124 (2002)
26. Satyanarayanan, M.: Pervasive computing: vision and challenges. IEEE Personal Communications 8, 10–17 (2002)
27. Portela, F., Santos, M.F., Silva, A.: Towards Pervasive and Intelligent Decision Support in Intensive Medicine – A Data Stream Mining Approach. In: AIM - Artificial Intelligence in Medicine (2013)
28. Guy, W.: ECDEU assessment manual for psychopharmacology. Rockville, Md (1976)
29. Guy, W., Rush, J., et al.: Clinical Global Impressions (CGI) Scale. Psychiatric Measures, APA (2000)