



## Artigo Completo

## Descoberta de regras de associação com algoritmo Tertius

## Discovery of association rules with Tertius algorithm

Norton Coelho Guimarães<sup>1</sup>, Rosângela Nunes<sup>2</sup><sup>1</sup>IF Goiano campus Morrinhos, BR 153 KM 633 Zona Rural, Morrinhos, Goiás, Brasil.<sup>2</sup>Universidade Federal de Goiás, CERCOMP, Goiânia, Goiás, Brasil. rdsnunes@gmail.com

\* Autor para correspondência: norton.guimaraes@ifgoiano.edu.br

## INFO ARTIGO

Histórico do artigo  
 Recebido: 24/09/2016  
 Aceito: 27/03/2017

**Palavras chaves:**

Mineração de Dados,  
 Regras de Associação,  
 Tertius.

**Keywords:**

Data Mining,  
 Associate Rules,  
 Tertius.

## RESUMO

O trabalho descreve os experimentos do processo de definição das regras de associação com base nas regras do desafio PhysioNet de 2012. O desafio disponibilizou uma base com 4000 pacientes com informações das últimas 48 horas de permanência em uma UTI. Na descoberta das regras de associação, agrupamos os registros por paciente e calculamos a média aritmética de cada atributo temporal. Para realizar os treinamentos utilizamos somente as primeiras 24 horas agrupadas de 6 em 6 horas. Utilizamos o algoritmo Tertius para gerar automaticamente as regras em conjunto com a ferramenta Weka.

## ABSTRACT

This paper describes experiments in the definition of association rules based on the rules of the challenge 2012 Physionet process. The challenge provided a base with 4000 patients with information of the last 48 hours of stay in the ICU. The discovery of association rules, we grouped the records by patient and calculated the arithmetic mean of each temporal attributes. To perform the training use only the first 24 hours grouped into six 6 hours. We use the algorithm Tertius to automatically generate rules together with the Weka tool.

## 1. Introdução

Uma Unidade de Tratamento Intensivo (UTI) é para pacientes que necessitam de tratamentos especiais, com uso de aparelhos, na maioria das vezes. Estes pacientes são monitorados frequentemente e são coletados diversos parâmetros fisiológicos para serem analisados posteriormente (XIA et al., 2012).

Atualmente, grupos de saúde estão focados em pesquisar novas técnicas para melhorar a eficácia do tratamento para os pacientes em estado crítico na UTI (BERA e NAYAK, 2012). Apesar destes esforços, o potencial das informações digitais não é plenamente utilizado (POLLARD et al., 2012).

Por décadas, o custo de permanência na UTI vem se elevando (XIA et al., 2012). Em 2005, o custo médio de um

paciente na UTI variava de 32 mil dólares à 42 mil dólares, para os que necessitavam de equipamentos para sobreviver, e para os pacientes que não necessitavam de aparelhos, constantemente, o custo médio se reduzia e ficava entre 12 mil dólares à 20 mil dólares (GOLDBERGER et al., 2000; XIA et al., 2012).

De um modo geral, para resolver o problema de mortalidade na UTI é utilizado a técnicas de regressão linear. Mas, podemos encontrar outros trabalhos que utilizem redes neurais, redes bayesianas, SVM e entre outras (XIA et al., 2012).

Este trabalho, é uma resposta para o desafio Computing in Cardiology/Physionet Challenge 2012 "Predicting Mortality of ICU Patients". O foco do desafio é desenvolver um método de predição da mortalidade hospitalar na UTI (GOLDBERGER et al., 2000).

Várias pesquisas têm aplicado diversas técnicas para encontrar soluções que possam resolver o problema de

mortalidade em UTI. Dentre estas técnicas, pode-se citar: Clusterização (XU et al., 2012), Lógica de Regressão (BERA e NAYAK, 2012; HAMILTON e HAMILTON, 2012), Redes Neurais (XIA et al., 2012), Redes Bayesianas (MACAS, 2012), SVM (BOSNJAK e MONTILLA, 2012), Motif Discovery (MCMILLAN, 2012) e entre outras.

Alguns dos estudos analisados propuseram que o próprio algoritmo manipule a falta de informação (JOHNSON et al., 2012). Outros, propuseram um método para manipular a falta de informação através da atribuição de valores. Em Lee (LEE et al., 2012), os valores faltantes foram preenchidos pela média considerando a idade e o gênero. Em Citi (CITI e BARBIERI, 2012), supôs que a falta de informação indicava uma situação estacionária, e propôs a replicação dos valores medidos anteriormente.

O objetivo, deste trabalho, é desenvolver um método para realizar a descoberta das regras de associação a partir dos registros disponibilizados.

Este trabalho, está organizado da seguinte forma. Na seção 3, descrevemos o desafio e suas etapas. Na seção 2, descrevemos os trabalhos relacionados. Na seção 4, descrevemos o método proposto para a descoberta das regras de associações. Na seção 5, descrevemos os experimentos e os resultados realizados durante a pesquisa. Por último, a conclusão é apresentada.

## 2. Material e Método

O PhysionNET/Computing Challenge é um desafio anual, iniciado em 1999, pelo National Institutes of Health e mantido pelo National Institute of Biomedical Imaging and Bioengineering (NIBIB) e National Institute of General Medical Sciences (NIGMS) com o intuito de desenvolver métodos para solução de temas específicos na área médica. Na edição de 2012, o foco foi o desenvolvimento de métodos para a predição de mortalidade de pacientes na UTI. Os dados usados para o desafio, consiste em 5 descrições gerais (veja detalhes na tabela 1) e 37 series temporais (veja detalhes na Tabela 3). Dados estes que foram coletados para às primeiras 48 horas iniciais desde a entrada do paciente na UTI e sua saída (veja detalhes na tabela 2). Foram disponibilizados registros de 12 mil pacientes adultos do banco de dados MIMIC II (SAEED et al., 2012).

As medições, das séries temporais, são registradas em ordem cronológica e podem ser registradas em intervalos regulares, variando de hora em hora, diariamente ou em intervalos irregulares, conforme necessário. Nem todas as séries temporais foram coletadas em todos os casos (SAEED et al., 2012).

Tabela 1: Tabela das descrições gerais (GOLDBERGER et. al., 2000)

Parâmetro	Descrição
<u>RecordID</u>	Identificação do paciente
<u>Age</u>	Idade do paciente
<u>Gender</u>	0: mulher ou 1 homem.
<u>Height</u>	Altura do paciente em cm
<u>ICUType</u>	Tipo da Série Temporal.
<u>Weight</u>	Peso do paciente em kg

Tabela 2: Tabela das descrições de saída do paciente (GOLDBERGER et. al., 2000)

Parâmetro	Descrição
<u>RecordID</u>	Identificação do paciente
<u>SAPS-I score</u>	Referência do estado psíquico
<u>SOFA score</u>	Referência da morte
<u>Length of stay</u>	Quantidade de dias de permanência no hospital
<u>Survival</u>	Quantidade dias de permanência na UTI
<u>In-hospital death</u>	0: Sobreviveu; 1: Morreu

Tabela 3: Tabela das séries temporais (GOLDBERGER et. al., 2000)

Tipo	Descrição	Medida
<i>Albumin</i>		(g/dL)
<i>ALP</i>	<i>Alkaline phosphatase</i>	(IU/L)
<i>ALT</i>	<i>Alanine transaminase</i>	(IU/L)
<i>AST</i>	<i>Aspartate transaminase</i>	(IU/L)
<i>Bilirubin</i>		(mg/dL)
<i>BUN</i>	<i>Blood urea nitrogen</i>	(mg/dL)
<i>Cholesterol</i>		(mg/dL)
<i>Creatinine</i>	<i>Serum creatinine</i>	(mg/dL)
<i>DiasABP</i>	<i>Invasive diastolic arterial blood pressure</i>	(mmHg)
<i>FiO2</i>	<i>Fractional inspired O2</i>	(0-1)
<i>GCS</i>	<i>Glasgow Coma Score</i>	(3-15)
<i>Glucose</i>	<i>Serum glucose</i>	(mg/dL)
<i>HCO3</i>	<i>Serum bicarbonate</i>	(mmol/L)
<i>HCT</i>	<i>Hematocrit</i>	(%)
<i>HR</i>	<i>Heart rate</i>	(bpm)
<i>K</i>	<i>Serum potassium</i>	(mEq/L)
<i>Lactate</i>		(mmol/L)
<i>Mg</i>	<i>Serum magnesium</i>	(mmol/L)
<i>MAP</i>	<i>Invasive mean arterial blood pressure</i>	(mmHg)
<i>MechVent</i>	<i>Mechanical ventilation respiration</i>	(0:false, or 1:true)
<i>Na</i>	<i>Serum sodium</i>	(mEq/L)
<i>NIDiasABP</i>	<i>Non-invasive diastolic arterial blood pressure</i>	(mmHg)
<i>NIMAP</i>	<i>Non-invasive mean arterial blood pressure</i>	(mmHg)
<i>NISysABP</i>	<i>Non-invasive systolic arterial blood pressure</i>	(mmHg)
<i>PaCO2</i>	<i>Partial pressure of arterial CO2</i>	(mmHg)
<i>PaO2</i>	<i>Partial pressure of arterial O2</i>	(mmHg)
<i>pH</i>	<i>Arterial pH</i>	(0-14)
<i>Platelets</i>		(cells/nL)
<i>RespRate</i>	<i>Respiration rate</i>	(bpm)
<i>SaO2</i>	<i>O2 saturation in hemoglobin</i>	(%)
<i>SysABP</i>	<i>Invasive systolic arterial blood pressure</i>	(mmHg)
<i>Temp</i>	<i>Temperature</i>	(°C)

<i>TropI</i>	<i>Troponin-I</i>	(mg/L)
<i>TropT</i>	<i>Troponin-T</i>	(mg/L)
<i>Urine</i>	<i>Urine output</i>	(mL)
<i>WBC</i>	<i>White blood cell count</i>	(cells/nL)
<i>Weight</i>		(kg)

O desafio de 2012, foi organizado em duas etapas: 1a - medição da performance do classificador binário; 2a - medição da performance do risco estimado. A pontuação da 1a etapa foi o valor baixo sensitivo e positiva para predição. A pontuação da 2a etapa foi o tamanho normalizado da estatística de Hosmer-Lemeshow (SAEED et. al., 2012).

Os dados utilizados na ICU 2012 foram extraídos do bando de dados clínico MIMIC II na versão 2.6 (SAEED et. al., 2012).

Foram selecionados 12 mil registros aleatoriamente de 12.753 registros com idade superior a 16 anos e com valores de entrada na UTI menor que 48 horas. Foram disponibilizados os registros em 3 agrupamentos (A, B e C) de 4000 pacientes (SAEED et. al., 2012).

No desafio de 2012, disponibilizaram o agrupamento A (treinamento) com as informações de estadia na UTI (treinamento), o agrupamento B (teste) para implementação do método de predição da mortalidade na UTI sem as informações de estadia na UTI e o agrupamento C somente para validar os resultados do método proposto, caso, a solução fosse selecionada para o final do desafio (SAEED et. al., 2012).

Utilizamos o conjunto de treinamento A. Este conjunto, consiste nos registros médicos de 4.000 pacientes, durante a estadia na UTI. Para cada paciente foram disponibilizados um arquivo do tipo ".txt" com as informações coletadas, conforme descrito na tabela 3, das primeiras 48 horas do paciente na UTI e as informações gerais do paciente, descrito nas tabelas 1 e 2. O nome do arquivo ".txt" é o identificador do paciente.

Após a análise dos arquivos, verificamos que não seria possível utiliza-los, na forma original, e escolhemos migrar os dados de cada arquivo em 03 tabelas (Paciente, Temporais e Outcomes) com relacionamentos, entre elas, através do ID do paciente.

Um dos problemas relatados no conjunto de dados disponibilizado foi a falta de informação em alguns dados dos pacientes. Para resolver esse problema atribuímos o valor nulo que posteriormente será discretizado como "?".

Para minimizar o ruído nos dados, resolvemos agrupar os dados pelo "id do paciente" e calcular a média aritmética de todas os atributos temporais.

Após a migração dos registros verificamos que a variável "id do paciente" não fornecia nenhuma informação relevante neste estudo, então, preferimos descartá-la. A ideia, inicial, é usar todas as variáveis restantes neste estudo e, deixar o algoritmo de associação tratar as possíveis redundâncias e/ou as variáveis não informativas.

Para definir as regras de associação utilizamos os registros das últimas 24 horas sendo agrupados de 6 em 6 horas e, geramos 4 grupos (G1, G2, G3 e G4) e 2 duas classes (F\_sim = Faleceu, F\_nao = Sobreviveu).

Para utilizar o algoritmo selecionado é necessário discretizar os dados (FLACH e LACHICHE, 1999). O ideal seria ter o apoio de um especialista nesta tarefa. Entretanto, por não termos um especialista participando deste trabalho, e devido ao tempo disponível para sua conclusão, optamos por utilizar somente os valores referenciados no material disponível no desafio para realizar a tarefa de discretização. Então, cada um dos valores associados aos registros dos pacientes foi mapeado para os valores discretos, conforme informações

apresentadas nas regras do desafio. Vejamos os detalhes na tabela 4

Tabela 4: Tabela das discretizações

Parâmetro	Descrição
Albumin, ALP, ALT, AST, Bilirubin, bun, Cholesterol, Creat, FiO2, Glucose, HCO, HCT, HR, K, Lactate, Mg, MAP, Na, NIMAP, PaCO2, PaO2, ph, Patelets, RespRate, SaO2, Temperatura, TroponinT, Urina, WBC, Weight	B: Baixo; N: Normal; A: Alto
DiasABP, NIDiasABP, NISysABP, SysABP	H; D; P; Stage_1; Stage_2; E
GCS	Severe; Moderate; Minor
MechVent	true; false
TroponinI	N: Normal; aN: Anormal

Por fim, utilizamos a ferramenta Weka (HALL et. al., 2009) juntamente com o algoritmo Tertius (FLACH e LACHICHE, 1999) na geração automática das regras de associação.

O algoritmo Tertius, faz o uso de uma abordagem heurística. Através de um algoritmo tipo "A\*", é realizada uma busca no espaço de possibilidades das regras de associação. A heurística utiliza uma distribuição Qui quadrado X2 que fornece uma regra  $A \Rightarrow B(\alpha, \beta)$  (FLACH e LACHICHE, 1999).

O objetivo do algoritmo consiste em aplicar um best-first-search, com a abordagem "A\*", encontrando as k hipóteses mais confirmadas, procurando regras com múltiplas condições, como o Apriori, e também inclui um refinamento de operadores não redundantes para anular as buscas desnecessárias. Os resultados são as regras estimadas, seguidas pelas observadas e por fim as confirmadas (FLACH e LACHICHE, 1999).

É com base nessas informações que um analista especializado pode realizar suas considerações e efetivar suas próximas estratégias. Ao final, o Tertius descreve a quantidade de hipóteses consideradas e a quantidade explorada (ANDRADE, 2007).

O algoritmo Tertius, demonstrou melhores resultados do que o algoritmo Predictive Apriori (SCHEFFER, 2005).

### 3 - Resultados e Discussão

Após a análise dos arquivos disponibilizados tivemos que migrar para um banco de dados relacional para facilitar a visualização das informações. Para auxiliar nessa migração foi desenvolvido uma aplicação Web<sup>1</sup> que gerou um arquivo ".arff"<sup>ii</sup>. Este arquivo, contém 4 grupos com os 37 atributos temporais, o atributo sexo e a classe (Faleceu, Sobreviveu). Vejamos no código 1 o exemplo dos cabeçalhos.

```
Código 1 Exemplo do cabeçalho do arquivo ARFF
1 @ATTRIBUTE Albumin_1 {B,N,A}
2 @ATTRIBUTE ALP_1 {B,N,A}
3 @ATTRIBUTE ALT_1 {B,N,A}
4 @ATTRIBUTE AST_1 {B,N,A}
```

5 ...  
 6 @ATTRIBUTE sexo {F,M}  
 7 @ATTRIBUTE class {F\_sim,F\_ao}

Realizamos diversos testes com o algoritmo Tertius que levou em média 2 horas de processamento. Durante os testes percebemos que o algoritmo utilizou um pico de 2,5Gb de RAM e um pico de 95% de processamento. A configuração da máquina utilizada para os experimentos foi um Intel Core i3 M370 2.40GHz, 4Gb de RAM, sistema operacional Windows 7 (64 bits), versão do Java 7.0 e versão do Weka 3.6.10.

Após a geração automática das regras pela ferramenta Weka com o uso do algoritmo Tertius foi necessário avaliar se as regras traziam algum conhecimento a respeito dos dados.

Diversas regras foram encontradas e os resultados devem ser interpretados de acordo com os aspectos procurados. Por não ter a avaliação de um especialista na área tivemos que fazer a análise com base nas referências disponíveis (GOLDBERGER et al., 2000). A seguir, serão apresentados os 18 melhores resultados encontrados. Vejamos na tabela 5 os detalhes dos resultados.

Na primeira regra, “HR\_2 = N and GCS\_4 = Minor”, representa que entre às 6 horas e 12 horas de internação na UTI o paciente tiver os batimentos cardíacos normais e entre às 18 horas e 24 horas estiver consciente, então, sobreviverá. Foram observados esta regra em 44,34% dos casos e confirmado em 20,45%.

A falta de um especialista dificultou a validação da solução proposta e, com isso, tivemos um grande esforço na análise, interpretação e preparação dos dados.

Durante a realização dos experimentos, nos deparamos com diversos problemas de recursos computacionais que impediram de utilizarmos o algoritmo PredictiveApriori como previsto anteriormente. O motivo de não utilizá-lo foi o alto consumo de memória que ultrapassou os 3Gb de memória RAM na máquina Virtual Java e, com isso, impossibilitou a obtenção dos resultados.

Chegamos a conclusão que não seria possível continuar os testes e resolvemos abandonar a utilização do algoritmo PredictiveApriori e passamos a utilizar o algoritmo Tertius disponível na ferramenta Weka e obtivemos êxito nos resultados.

Podemos perceber que a descoberta das regras de associação, é um trabalho bastante complexo, se não tivermos a presença de um especialista, para avaliá-las.

Para os trabalhos futuros, pretendemos retomar os estudos com o algoritmo PredictiveApriori, e comparar com os resultados do algoritmo Tertius.

## 6. References

ANDRADE, A. A. d. Desenvolvimento de sistema especialista com operacionalidade de aprendizado para operar em tempo real com sistemas industriais automatizados. PhD thesis, Escola Politécnica - USP, São Paulo, Brasil, abr 2007.

BERA, D.; NAYAK, M. Mortality risk assessment for icu patients using logistic regression. In: Computing in Cardiology (CinC), 2012, p. 493–496, 2012.

BOSNJAK, A.; MONTILLA, G. Predicting mortality of ICU patients using statistics of physiological variables and support vector machines. In: Computing in Cardiology (CinC), 2012, p. 481–484, 2012.

CITI, L.; BARBIERI, R. Physionet 2012 challenge: Predicting mortality of icu patients using a cascaded svm-glm paradigm. In: Computing in Cardiology (CinC), 2012, p. 257–260, 2012.

FLACH, P. A.; LACHICHE, N. Confirmation-guided discovery of first-order rules with tertius. Machine Learning, 42:61–95, 1999.

GOLDBERGER, A. L.; AMARAL, L. A. N.; GLASS, L.; HAUSDORFF, J. M.; IVANOV, P. C.; MARK, R. G.; MIETUS, J. E.; MOODY, G. B.; PENG, C.-K.; STANLEY, H. E. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation, 101(23):e215–e220, 2000 (June 13). Circulation Electronic Pages: <http://circ.ahajournals.org/cgi/content/full/101/23/e215> PMID:1085218; doi: 10.1161/01.CIR.101.23.e215.

HALL, M.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, I. H. The weka data mining software: an update. SIGKDD Explor. Newsl., 11(1):10–18, Nov. 2009.

HAMILTON, S.; HAMILTON, J. Predicting in-hospital-death and mortality percentage using logistic regression. In: Computing in Cardiology (CinC), 2012, p. 489–492, 2012.

No	Esperados	Observados	Confirmados	Regras (Sobreviveu)
1	15,70	44,34	20,45	HR_2 = N and GCS_4 = Minor
2	15,60	42,05	18,76	HR_2 = N and GCS_4 = Minor and Urina_4 = B
3	15,47	42,05	18,94	HR_2 = N and Urina_3 = B and GCS_4 = Minor
4	15,42	54,05	28,89	Urina_3 = B and GCS_4 = Minor
5	15,27	42,17	19,32	HR_2=N and Urina_2=B and GCS_4=Minor
6	15,18	43,41	20,45	HR_1 = N and GCS_4 = Minor
7	15,17	44,31	21,20	GCS_4 = Minor and HR_4 = N
8	15,17	54,02	29,26	GCS_4 = Minor and Urina_4 = B
9	15,12	42,35	19,69	Urina_3=B and GCS_4=Minor and HR_4=N
10	15,09	40,42	18,19	HR_2=N and HR_3=N and GCS_4=Minor
11	15,00	41,21	18,94	HR_1 = N and GCS_4 = Minor and Urina_4 = B
12	15,00	54,14	29,64	Urina_2 = B and GCS_4 = Minor
13	14,97	52,75	28,51	Urina_3 = B and GCS_4 = Minor and Urina_4 = B
14	14,93	42,02	19,69	GCS_4 = Minor and HR_4 = N and Urina_4 = B
15	14,90	41,27	19,13	HR_1 = N and Urina_3 = B and GCS_4 = Minor
16	14,76	42,17	20,07	Urina_2 = B and GCS_4 = Minor and HR_4 = N
17	14,71	41,15	19,32	Urina_1 = B and HR_2 = N and GCS_4 = Minor
18	14,64	41,27	19,51	HR_1 = N and Urina_2 = B and GCS_4 = Minor

Na décima oitava regra, “HR\_1 = N and Urina\_2 = B and GCS\_4 = Minor”, representa que entre às 0 hora e 6 horas de internação na UTI o paciente tiver os batimentos cardíacos normais e entre às 6 horas e 12 horas estiver com a Ureia baixa e entre às 18 horas e 24 horas estiver consciente, então, sobreviverá. Foram observados esta regra em 41,27% dos casos e confirmado em 19,51%.

Podemos concluir que caso o paciente não apresente as regras nas suas medições, provavelmente, o paciente não sobreviverá.

O objetivo do nosso trabalho não é descrever todas as regras. Podemos descrevê-las, com base nas explicações anteriores aplicando as informações da tabela 3, e as descrições detalhadas de cada atributo temporal disponível no desafio (GOLDBERGER et al., 2000).

## 4. Conclusão

Durante os estudos nos deparamos com muitos artigos referentes ao tema da Mortalidade na UTI.

JOHNSON, A.; DUNKLEY, N.; MAYAUD, L.; TSANAS, A.; Kramer, A.; Clifford, G. Patient specific predictions in the intensive care unit using a bayesian ensemble. In: Computing in Cardiology (CinC), 2012, p. 249–252, 2012.

LEE, C.; ARZENO, N.; HO, J.; VIKALO, H.; GHOSH, J. An imputation-enhanced algorithm for icu mortality prediction. In: Computing in Cardiology (CinC), 2012, p. 253–256, 2012.

MACAS, M.; KUZILEK, J.; ODSTRCILIK, T.; HUPTYCH, M. Linear bayes classification for mortality prediction. In: Computing in Cardiology (CinC), 2012, p. 473–476, 2012.

MCMILLAN, S.; CHIA, C.-C.; VAN ESBROECK, A.; RUBINFELD, I.; SYED, Z. Icu mortality prediction using time series motifs. In: Computing in Cardiology (CinC), 2012, p. 265–268, 2012.

POLLARD, T.; HARRA, L.; WILLIAMS, D.; HARRIS, S.; MARTINEZ, D.; FONG, K. 2012 physionet challenge: An artificial neural network to predict mortality in icu patients and application of solar physics analysis methods. In: Computing in Cardiology (CinC), 2012, p. 485–488, 2012.

SAEED, M.; VILLARROEL, M.; REISNER, A.; CLIFFORD, G.; LEHMAN, L.; MOODLY, G.; HELDT, T.; KYAW, T.; MOODY, B.; MARK, R. Multiparameter intelligent monitoring in intensive care ii (mimic ii): A public-access intensive care unit database. Critical Care Medicine, 2011.

SCHEFFER, T. Finding association rules that trade support optimally against confidence. Intell. Data Anal., 9(4):381–395, July 2005.

XIA, H.; DALEY, B.; PETRIE, A.; ZHAO, X. A neural network model for mortality prediction in icu. In: Computing in Cardiology (CinC), 2012, p. 261–264, 2012.

XU, J.; LI, D.; ZHANG, Y.; DJULOVIC, A.; LI, Y.; ZENG, Y. Cinc challenge: Cluster analysis of multi-granular time-series data for mortality rate prediction. In: Computing in Cardiology (CinC), 2012, p. 497–500, 2012.

---

<sup>i</sup> Utilizamos o software XAMPP v.3.1.0 com MySql 3.5.2.2 e PHP 5.2

<sup>ii</sup> ARFF é um arquivo de texto ASCII que descreve uma lista de instâncias de partilha de um conjunto de atributos utilizado no Weka