Classificação automática de dados para descoberta de conhecimento: um estudo de caso para classificação de risco na área da saúde

João Paulo Funchal¹ Carlos Alberto Madsen² Diana Francisca Adamatti¹

Resumo: O presente artigo tem como objetivo apresentar um estudo da utilização de técnicas de mineração de dados no contexto da classificação de risco em unidades de pronto atendimento. Foi utilizada a técnica conhecida como árvores de decisão, tendo em vista sua ampla utilização em problemas de classificação e, também, com o intuito de tornar-se uma ferramenta essencial para um sistema de apoio à decisão mais alinhado com o conhecimento do profissional de saúde. Todavia, é imprescindível ressaltar que a tomada de decisão do risco de vida do paciente vem do conhecimento técnico e da experiência desse profissional, que normalmente vão além do conhecimento teórico.

Palavras-chave: Classificação de riscos. Mineração de dados. Sistemas de apoio à decisão.

Abstract: This paper aims to present a study of the use of data mining techniques in the context of risk classification in emergency units. The technique known as decision trees, in view of its wide use in classification problems was used. And also in order to become an essential tool for a support system more aligned with the knowledge of the health professional decision. Stressing that the decision-making risk the patient's life comes from the knowledge and experience of this professional, it usually goes beyond theoretical knowledge.

Keywords: Data mining. Decision support systems. Risk classification.

1 Introdução

A utilização de inteligência artificial, mais especificamente de sistemas especialistas ou de apoio à decisão, na área de saúde não é recente, na medida em que, ao longo dos anos, diversos trabalhos abordaram esse tema. Todavia, esses trabalhos normalmente focam na questão do diagnóstico, tanto no âmbito da medicina como no da enfermagem [1, 2, 3, 4, 5]. No contexto do pronto atendimento, a etapa anterior ao diagnóstico, ou seja, o processo de acolhimento e classificação de risco dos pacientes, não tem sido abordada em trabalhos científicos que empreguem técnicas de mineração de dados e inteligência artificial.

Ao encontro desse fato, a técnica conhecida como árvore de decisão se destaca por ser amplamente utilizada em sistemas de tomada de decisão aplicados à saúde, tendo em vista o seu rápido aprendizado de conceitos e a sua facilidade de implementação [6, 7, 8].

http://dx.doi.org/10.5335/rbca.2015.4637

¹ Centro de Ciências Computacionais – Universidade Federal do Rio Grande – C3/FURG – Rio Grande (RS) - Brasil {funchaljoaopaulo,dianaada@gmail.com}

² Núcleo de Tecnologia da Informação - Universidade Federal do Rio Grande - NTI/FURG - Rio Grande (RS) - Brasil {carlos.madsen@gmail.com}

Atualmente, ao ingressar no pronto atendimento do Hospital Universitário Dr. Miguel Riet Corrêa Jr., o paciente passa por um processo de triagem de risco, no qual, por meio de um sistema informatizado, o profissional de saúde define qual a prioridade para o seu atendimento. Esse sistema tem por objetivo registrar os dados vitais e sintomas dessa pessoa, deixando totalmente a cargo de quem o acolhe definir essa prioridade. Assim sendo, o presente trabalho tem por objetivo melhorar a avaliação dos dados inseridos, diminuindo as etapas para a classificação e garantindo ao profissional de saúde um apoio na sua tomada de decisão.

Este artigo está dividido em cinco seções. Na seção dois, são apresentados os conceitos sobre a classificação de risco, a forma como é utilizado o sistema de classificação e suas particularidades. A seção três trata brevemente de árvores de decisão e sua aplicação na classificação. Já na seção quatro, são mostrados o experimento e os resultados obtidos da mineração dos dados do sistema. Por fim, a seção cinco apresenta as conclusões encontradas até o momento da classificação dos dados.

2 Classificação de risco

A classificação de risco, ou triagem, tem por objetivo a regulação da demanda do atendimento nos serviços de pronto atendimento, visando identificar o atendimento médico imediato e o mediato, organizando, assim, o fluxo de pacientes que procuram essas unidades [9, 10]. Apesar do excesso de demanda que sofre esse tipo de unidade hospitalar, utilizando desse processo de triagem são minimizados os riscos para os pacientes que as procuram [11].

A implantação da classificação de risco em todas as unidades de pronto atendimento é proposta pela Portaria 2.048/2009 do Ministério da Saúde, devendo esse processo de classificação ser realizado por profissional de saúde, de nível superior, via de regra, um enfermeiro, com treinamento específico e se valendo de protocolos de acolhimento previamente estabelecidos [12].

No protocolo proposto pelo Sistema Único de Saúde (SUS), utilizado neste trabalho, mediante o julgamento crítico e a experiência do enfermeiro, o paciente pode ser classificado nas seguintes cores [10]:

- VERMELHO: emergência, o atendimento será realizado imediatamente na sala de emergências;
- AMARELO: urgência, terá o atendimento priorizado sobre os pacientes classificados como VERDE, no consultório ou leito da sala de observação;
- VERDE: sem risco de morte imediato, tendo o atendimento realizado depois que todos os pacientes classificados como VERMELHO e AMARELO forem atendidos;
- AZUL: quadro crônico sem sofrimento agudo ou caso social, deverá ser preferencialmente encaminhado para atendimento em Unidade Básica de Saúde ou atendido pelo Serviço Social.

Por fim, é importante frisar que essa prática não se caracteriza pelo diagnóstico da doença, mas sim pela determinação da prioridade de atendimento, de acordo com a gravidade do estado do paciente [9].

2.1 Sistema

O presente sistema de classificação de risco foi implantado em julho de 2012 no Serviço de Pronto Atendimento (SAP) do Hospital Universitário Dr. Miguel Riet Corrêa Jr., sendo inicialmente baseado no protocolo de acolhimento do SUS [10]. Todavia, o sistema permite que os enfermeiros responsáveis pelo serviço de pronto atendimento cadastrem novos sintomas, os relacionem a qualificadores e, por fim, definam a sua classificação, permitindo, assim, que este serviço se adapte a realidade local.

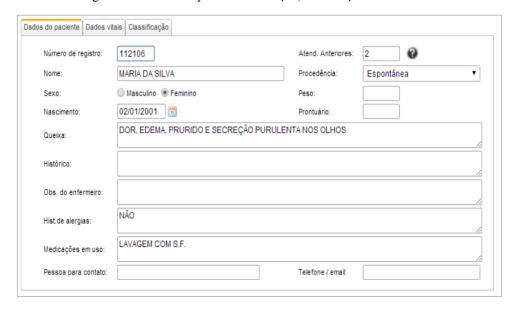
Por exemplo, atualmente o sintoma diabetes está configurado de acordo com a Tabela 1:

Tabela 1: Exemplo de qualificadores e seus classificadores para o sintoma de diabetes.

| Sintoma | Qualificador | Classificador |
|----------|--|-----------------|
| Diabetes | Sudorese profusa (hipoglicemia) | VERMELHO |
| | Alteração do estado mental (letargia, confusão mental, agitação, coma) | VERMELHO |
| | Dados vitais alterados | VERMELHO |
| | Glicemia >250 mg/dl e sinais de desidratação | AMARELO |
| | Dados vitais normais | AMARELO |
| | Glicemia > 250mg/dl e assintomático | VERDE |
| | Glicemia <= 250mg/dl e assintomático | AZUL |
| | História de diabetes e precisando de medicação (receita) | AZUL |

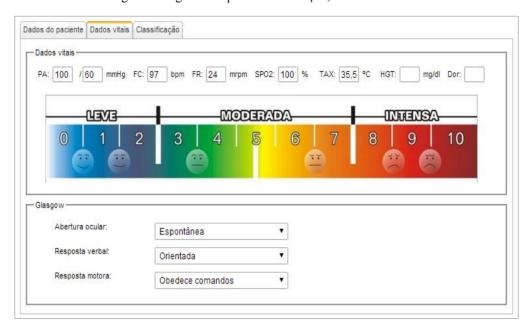
Nesse sistema, a interface de classificação de risco é dividida em quatro etapas distintas. Inicialmente, o enfermeiro registra os dados básicos do paciente, conforme ilustrado na Figura 1.

Figura 1: Primeira etapa da classificação, informações básicas.



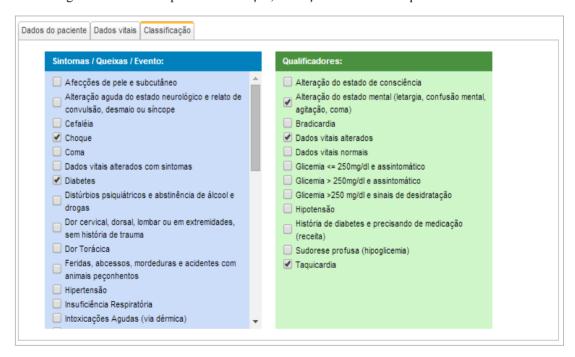
Em seguida, de acordo com a Figura 2, são informados ao sistema os dados vitais do paciente, sua escala de dor e de coma (escala de Glasgow).

Figura 2: Segunda etapa da classificação, dados vitais.



Na sequência, como apresentado na Figura 3, o profissional pode escolher os sintomas que identifica no paciente (quadro azul da esquerda) e seus respectivos qualificadores (quadro verde da direita). Destaca-se a possibilidade de se marcar diversos sintomas e qualificadores.

Figura 3: Terceira etapa da classificação, definição de sintomas e qualificadores.



Por fim, como ilustrado na Figura 4, o sistema contabiliza o número de classificadores relacionados aos qualificadores selecionados (número entre parênteses) e, dado o maior número, sugere uma classificação. No entanto, a decisão final fica a cargo do profissional de saúde, que tem a possibilidade de escolher a classificação que considere relevante para o caso.



Figura 4: Quarta etapa da classificação, definição do classificador e registro.

O sistema de classificação de riscos do SAP é utilizado desde 27/07/2012, já tendo sido registrados 74.075 atendimentos, com a média de um atendimento a cada 15 minutos.

3 Árvores de decisão

Árvores de decisão são algoritmos utilizados para a classificação e a previsão de dados. Os algoritmos de indução a árvores de decisão se baseiam na estratégia de divisão e conquista, que consiste na recursiva divisão do problema em problemas menores, até que a solução dos subproblemas seja o mais simples possível [13]. Sua construção, basicamente, utiliza um conjunto de treinamento formado por entradas e saídas de informação. Os algoritmos apresentam uma estrutura não linear e um número finito de elementos e nodos; sua representação gráfica torna mais simples sua prática, podendo ser seguido como um passo a passo, começando pelo nó raiz (principal) até seus nós folhas, onde cada nó da árvore é um atributo e para cada atributo uma decisão será tomada [14].

As árvores de decisão estão entre os mais populares algoritmos de inferência e têm sido aplicadas em várias áreas, como diagnóstico médico e risco de crédito, e delas se pode extrair regras do tipo "se então" que são facilmente compreendidas. Seu principal uso na mineração de dados se dá pela facilidade de entender e interpretar. É um algoritmo robusto e tem bom desempenho com grandes quantidades de informação em pouco tempo [13, 14].

4 Experimentos e resultados

Neste trabalho, foram analisados os dados vitais de pacientes que tiveram sua classificação de risco registrada nos anos de 2012 e 2013.

A ferramenta escolhida para o processamento dos dados é o Weka, um software que possui uma coleção de algoritmos para tarefas de mineração, além de ferramentas para pré-processamento, classificação, agrupamentos, entre outras [15]. Para o experimento, foram escolhidos os atributos utilizados em mais de 60% das classificações registradas no sistema, sendo eles: sexo, idade, pressão arterial sistólica (PAS), pressão arterial diastólica (PAD), frequência cardíaca (FC), frequência respiratória (FR), saturação parcial de oxigênio (SPO2), glasgow ocular (abertura ocular), glasgow verbal (resposta verbal), glasgow motora (resposta motora), temperatura e a cor em que o paciente foi classificado.

Os resultados apresentados na Figura 5 são de todos os registros feitos entre 2012 e 2013, em um total de 55.167 pacientes, com uma taxa de acerto de 60,22%. A árvore foi configurada da seguinte maneira: mínimo de duzentas instâncias por folha com 0.25 de fator de confidência.

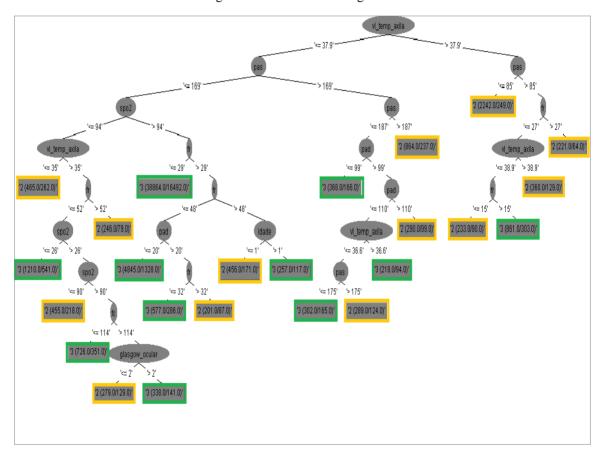


Figura 5: Árvore de decisão gerada.

Pode-se perceber que a árvore de decisão foi montada utilizando a temperatura como nó raiz e encontrando associações mais rapidamente para os casos à esquerda da árvore.

A matriz de confusão gerada também nos fornece informações importantes, como ilustra a Figura 6. É importante ressaltar que a diagonal principal da matriz de confusão deve sempre conter o maior valor da linha e que a taxa de erro ainda é alta devido à grande variação nos dados. Podemos concluir, pela matriz, que essa configuração na árvore de decisão apresenta resultados inconclusos. Sendo assim, ainda se fazem necessárias

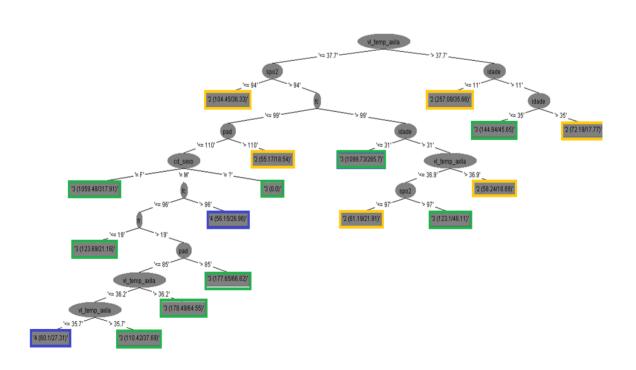
algumas modificações, buscando aperfeiçoar a configuração da árvore de decisão, para então obtermos melhores resultados.

Figura 6: Matriz de confusão gerada.

| 0.952 0 0.37 | 0.795 0 0.046 | Precision 0.589 0 0.706 | 0.952 0 | 0.727 | 0.608 | 3 |
|--------------------|-------------------------|---|--|--|--|---------------------------|
| 0 0.37 | 0 0.046 | 0 | 0 | | | |
| 0.37 | 0.046 | | | 0 | 0.627 | 4 |
| | | 0.706 | 0.00 | | -1027 | 4 |
| 0 | | | 0.37 | 0.486 | 0.703 | 2 |
| _ | 0 | 0 | 0 | 0 | 0.752 | 1 |
| 0.603 | 0.443 | 0.482 | 0.603 | 0.507 | 0.635 | |
| | | | | | | |
| | | | | | | |
| 1442 0 | a = | = 3 | | | | |
| 218 0 | b = | = 4 | | | | |
| 4663 0 | c = | = 2 | | | | |
| 278 0 | d = | = 1 | | | | |
| | c d 1442 0 218 0 4663 0 | C d < cla 1442 0 a = 218 0 b = 4663 0 c = | Matrix === c d < classified as 1442 0 a = 3 218 0 b = 4 4663 0 c = 2 | Matrix === c d < classified as 1442 0 a = 3 218 0 b = 4 4663 0 c = 2 | Matrix === c d < classified as 1442 0 a = 3 218 0 b = 4 4663 0 c = 2 | c d < classified as 1442 |

Contudo, também aplicamos a mesma configuração com registros separados por enfermeiros que os classificaram e obtivemos resultados interessantes, como mostra a Figura 7.

Figura 7: Árvore de decisão gerada.



A Figura 7 representa a árvore de decisão de apenas um enfermeiro, com um total 3.742 pacientes registrados pelo profissional. Podemos perceber que a árvore apresenta uma estrutura muita parecida com a da Figura 5; porém, com um melhor balanceamento e uma melhor distribuição dos resultados em cada nó, sua taxa de acerto foi de 71,5%.

Figura 8: Matriz de confusão gerada.

| | | TI | Rate | FP Rate | Precision | Recall | F-Measure | ROC Area | Class |
|---------------|----------------------------|--------------------------|-----------------------|------------------|-----------|--------|-----------|----------|-------|
| | | | 0.243 | 0.022 | 0.548 | 0.243 | 0.336 | 0.816 | 4 |
| | | | 0.926 | 0.549 | 0.728 | 0.926 | 0.815 | 0.743 | 3 |
| | | | 0.496 | 0.054 | 0.777 | 0.496 | 0.606 | 0.793 | 2 |
| | | | 0 | 0 | 0 | 0 | 0 | 0.872 | 1 |
| eigh | ted Av | ď | 0.729 | 0.354 | 0.716 | 0.729 | 0.701 | 0.766 | |
| == C | | - | rix === | | 0.710 | 0.725 | 0.701 | 0.766 | |
| | onfusi | on Mat | rix === | | | 0.723 | 0.701 | 0.766 | |
| a | onfusi b | on Mat | rix === | | | 0.729 | 0.701 | 0.766 | |
| a 91 | onfusi b 273 | on Mat c 11 | crix ==== d < 0 | : (classif | | 0.729 | 0.761 | 0.766 | |
| a 91 48 | onfusi b 273 2127 | on Mat c 11 123 | d < | classif a = 4 | | 0.723 | 0.761 | 0.766 | |

Também podemos perceber, na Figura 8, que a matriz de confusão fica com uma maior taxa de acerto, o que pode ser um fato relevante, uma vez que, minerando os dados separados por enfermeiros, a árvore encontra um padrão mais rapidamente e obtêm-se melhores resultados.

Observando que tínhamos uma diferença significativa em relação à classificação dos pacientes, cuja maioria era classificada como prioridade(s) "verde" e/ou "azul", optamos por balancear esses registros, selecionando o mesmo número de paciente classificados em cada cor. Obtivemos o seguinte resultado, como nos mostra a Figura 9, em um total de 2.295 registros com uma taxa de acerto de 59,30%.

1 (25.06.01)

1 (25.06.01)

1 (25.06.01)

1 (25.06.01)

1 (25.06.01)

1 (25.06.01)

1 (25.06.01)

1 (25.06.01)

1 (25.06.01)

1 (25.06.01)

1 (25.06.01)

1 (25.06.01)

1 (25.06.01)

1 (25.06.01)

1 (25.06.01)

1 (25.06.01)

1 (25.06.01)

1 (25.06.01)

1 (25.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1 (26.06.01)

1

Figura 9: Árvore de decisão gerada.

Figura 10: Matriz de confusão gerada com registros balanceados.

| === I | Detai | iled | Accı | ıra | acy I | Ву С | lass === | = | | | | | | | |
|-------|-------|------|------|-----|-------|------|----------|-------|------|-------|----|----------|---|----------|-------|
| | | | T | P F | Rate | Ε | P Rate | Preci | sion | Recal | 11 | F-Measur | e | ROC Area | Class |
| | | | | 0. | 564 | | 0.113 | 0. | 619 | 0.5 | 64 | 0.59 | | 0.764 | 1 |
| | | | | 0. | .74 | | 0.31 | 0. | 445 | 0.7 | 4 | 0.556 | | 0.767 | 4 |
| | | | | 0. | 445 | | 0.097 | 0. | 606 | 0.4 | 45 | 0.513 | | 0.734 | 2 |
| | | | | 0. | 314 | | 0.126 | 0. | 456 | 0.3 | 14 | 0.372 | | 0.685 | 3 |
| Weigh | nted | Avg. | | 0. | .515 | | 0.162 | 0. | 531 | 0.5 | 15 | 0.507 | | 0.738 | |
| === (| | | | | | | ssified | 20 | | | | | | | |
| | 119 | | | | | | | as | | | | | | | |
| | 427 | | | | | | | | | | | | | | |
| | 141 | | | | | | | | | | | | | | |
| | 273 | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | |

5 Conclusão

A análise da classificação de risco é muito importante para a otimização do sistema de saúde, pois, além de agilizar os processos, pode tornar mais justo e mais eficiente o atendimento ao paciente. Portanto, não foram considerados, neste estudo, os sintomas e os respectivos qualificadores que foram selecionados, tendo em vista que o objetivo aqui consistiu em verificar se uma árvore de decisão teria a capacidade de mapear o conhecimento do profissional de enfermagem, utilizando somente os dados vitais do paciente. Percebe-se que ainda é imprescindível um estudo mais aprofundado do problema e da configuração da árvore de decisão, para a obtenção de melhores resultados. Não podemos deixar de considerar a falta de um padrão dos enfermeiros para classificação dos pacientes, tendo-se notado, nos experimentos, que essa lacuna influencia muito nos resultados, tornando mais difícil para a árvore de decisão encontrar associações.

Como trabalhos futuros, pretendemos melhorar a análise dos dados, bem como realizar alguns filtros nos dados, de forma a obter resultados mais precisos. Também pretendemos utilizar outras técnicas de mineração de dados, como agrupamentos (*clustering*) e redes neurais.

Referências

- [1] Souza A. R, Talon A. F., *Inteligência Artificial Aplicada à Medicina*. In: Cadernos de Estudos Tecnológicos, 2013; v. 1; n. 1; p. 59-72; ISSN: 2238-3948.
- [2] Farias E. P., Sassi R. J., Soares E. D. *Uso de Sistema Especialista no Auxílio ao Diagnóstico por meio de escalas de apoio à Enfermagem*. In: Simpósio Internacional de Inovação e Sustentabilidade, 2013, Uninove, São Paulo, SP.
- [3] Teixeira M. J. C. Soares T., Ferreira A., Pinto J.; Os contributos dos sistemas de apoio à decisão para a prática de enfermagem, In: Journal of Health Informatics, 2012, v. 4, n.2, p. 59-63; ISSN 2175-4411.
- [4] Sellmer D, Carvalho C. M. G., Carvalho D. R., Malucelli A. *Sistema especialista para apoiar a decisão na terapia tópica de úlceras venosas*; In: Revista Gaúcha de Enfermagem, 2013; v. 34, n. 2, p. 154-162, ISSN 0102-6933.
- [5] Lichtenstein F., Tavares A., Pisa I., Sigulem D., Sistemas de apoio à decisão baseados em diretrizes interpretadas por computador: um breve histórico e outros tópicos, In: Journal of Health Informatics, 2011, v. 3, n. 4, p. 164-169, ISSN 2175-4411.
- [6] Medeiros A. R. M., Araújo Y. B., Vianna R. P. T., Moraes R. M., *Modelo de suporte à decisão aplicado à identificação de indivíduos não aderentes ao tratamento anti-hipertensivo*, 2014, SAÚDE em DEBATE, v. 38, n. 100, p. 104-118, JAN-MAR, ISSN 0103-1104.
- [7] Tenório j. M, et al. Experiências internacionais da aplicação de sistemas de apoio à decisão clínica em gastroenterologia, In: Journal of Health Informatics, 2011, v. 3, n.1, p. 27-31; ISSN 2175-4411..
- [8] Chaves D. B R., Árvores de Decisão para inferência de Desobstrução ineficaz de vias aéres e Padrão respiratório ineficaz de Crianças com Infecção Respiratória Aguda. 2011, Universidade Federal do Ceará.
- [9] Servin S. C. N, et al; Protocolo de acolhimento e classificação de risco, Prefeitura de São Luís, Maranhão 2014. Disponível em: http://bvsms.saude.gov.br/bvs/publicacoes/protocolo_acolhimento_classificacao_risco.pdf. Acesso em 18 ago. de 2014.

- [10] Santos, E. A, et al; *Acolhimento e Classificação de Risco*, Prefeitura de Belo Horizonte, Minas Gerais 2014. Disponível em: http://www.pbh.gov.br/smsa/biblioteca/protocolos/AcolhimentoClassificacaodeRiscodasUpasdeBH.pdf. Acesso em 18 ago. de 2014.
- [11] Albino R. A, Grosseman S. Riggenach V., *Classificação de risco: Uma necessidade inadiável em um serviço de emergência de qualidade* In: Arquivos Catarinenses de Medicina, 2007. v. 36. n. 4 p. 70-75; ISSN 1806-4280.
- [12] Ministério da Saúde; Portaria 2048 de 2002. Disponível em: http://dtr2001.saude.gov.br/sas/PORTARIAS/Port2002/Gm/GM-2048.htm Acesso 18 agosto de 2014.
- [13] Von Zuben, Fernando J: Sistemas Baseados em Regras e Árvores de Decisão, Unicamp São Paulo. Disponivel em: <ftp://ftp.dca.fee.unicamp.br/pub/docs/vonzuben/ea072_2s11/topico6_EA072_2s11.pdf>. Acesso em 20 ago. de 2014.
- [14] Halmenschlager, Carine: *Um algoritmo para indução de árvores e regras de decisão*, URGS Porto Alegre 2002. Disponível em: http://www.lume.ufrgs.br/bitstream/handle/10183/2755/000325797.pdf?sequence=1. Acesso em 18 ago. de 2014.
- [15] Weka3: *Data Mining Software in Java*. Weka 2014. Disponível em: http://www.cs.waikato.ac.nz/ml/weka/>. Acesso em 18 ago. de 2014.