






## KNOWLEDGE DISCOVERY IN DATABASES PARA APOIO À DECISÃO EM ATIVIDADES VINCULADAS À ENFERMAGEM

### KNOWLEDGE DISCOVERY IN DATABASES FOR THE SUPPORT TO DECISION-MAKING IN NURSING-RELATED ACTIVITIES

### KNOWLEDGE DISCOVERY IN DATABASES PARA APOYAR LA DECISIÓN EN ACTIVIDADES RELACIONADAS CON LA ENFERMERÍA

 Ana Paula Sartorelli<sup>1</sup>  
 Carlyne Lopata<sup>1</sup>  
 Denilsen Carvalho Gomes<sup>1</sup>  
 Marcia Regina Cubas<sup>1</sup>  
 Deborah Ribeiro Carvalho<sup>1</sup>

<sup>1</sup>Pontifícia Universidade Católica - PUC, Tecnologia em Saúde. Curitiba, PR – Brasil.

Autor Correspondente: Márcia Regina Cubas  
E-mail: mcubas@pucpr.br

#### Contribuições dos autores:



Aquisição de Financiamento: Ana P. Sartorelli, Carlyne Lopata, Marcia R. Cubas, Deborah R. Carvalho; Coleta de Dados: Ana P. Sartorelli, Carlyne Lopata; Conceitualização: Ana P. Sartorelli, Carlyne Lopata, Denilsen C. Gomes, Marcia R. Cubas, Deborah R. Carvalho; Gerenciamento de Recursos: Ana P. Sartorelli, Carlyne Lopata, Denilsen C. Gomes; Gerenciamento do Projeto: Ana P. Sartorelli, Carlyne Lopata, Denilsen C. Gomes, Marcia R. Cubas, Deborah R. Carvalho; Investigação: Ana P. Sartorelli, Carlyne Lopata; Metodologia: Ana P. Sartorelli, Carlyne Lopata, Denilsen C. Gomes, Marcia R. Cubas, Deborah R. Carvalho; Redação - Preparação do Original: Ana P. Sartorelli, Carlyne Lopata, Denilsen C. Gomes, Deborah R. Carvalho; Redação - Revisão e Edição: Ana P. Sartorelli, Carlyne Lopata, Denilsen C. Gomes; Supervisão: Marcia R. Cubas, Deborah R. Carvalho; Validação: Ana P. Sartorelli, Carlyne Lopata, Denilsen C. Gomes, Marcia R. Cubas, Deborah R. Carvalho; Visualização: Ana P. Sartorelli, Carlyne Lopata, Marcia R. Cubas, Deborah R. Carvalho.

Fomento: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES. Código de Financiamento 001.

Submetido em: 19/09/2020

Aprovado em: 12/02/2021

#### Editores Responsáveis:

 Kênia Lara Silva  
 Tânia Couto Machado Chianca

## RESUMO

**Objetivo:** relatar experiências de estratégias do uso da mineração de dados em dois cenários de práticas de Enfermagem. **Descrição da experiência:** em ambas as experiências foi utilizado o algoritmo Apriori para descoberta de regras de associação e identificado as situações de exceção. A primeira experiência utilizou dados provenientes de óbitos infantis da região metropolitana de Curitiba - PR. Na segunda experiência utilizaram-se prontuários de pacientes atendidos por enfermeiros na classificação de risco em um hospital particular de Curitiba - PR. O primeiro estudo identificou 374 regras gerais e o segundo, 108 regras gerais, ambos com suas respectivas regras de exceção. **Conclusão:** a aplicação do Knowledge Discovery in Databases pode ser demonstrada e efetivada em dois cenários distintos, a fim de contribuir para a tomada de decisão pelo gestor. Espera-se que o relato reforce a importância do ensino da informática em Enfermagem como ferramenta de apoio à decisão.

**Palavras-chave:** Mineração de Dados; Bases de Dados Bibliográficas; Informática em Enfermagem; Educação em Enfermagem; Tomada de Decisões.

## ABSTRACT

**Objective:** to report experiences of data mining use strategies in two Nursing practice settings. **Description of the experience:** in both experiences, the Apriori algorithm was used to discover association rules and to identify exception situations. The first experience used data from infant deaths in the metropolitan region of Curitiba, PR. In the second experience, medical records of patients assisted by nurses were used in the risk classification at a private hospital in Curitiba -PR. The first study identified 374 general rules and the second, 108, both with their respective exception rules. **Conclusion:** the application of Knowledge Discovery in Databases can be demonstrated and carried out in two different settings, in order to contribute to decision-making by the manager. It is expected that the report reinforces the importance of teaching Nursing Informatics as a decision-making support tool.

**Keywords:** Data Mining; Databases, Bibliographic; Nursing Informatics; Education, Nursing; Decision Making.

## RESUMEN

**Objetivo:** reportar experiencias de estrategias de uso de minería de datos en dos escenarios de prácticas de enfermería. **Descripción del experimento:** en ambos experimentos se utilizó el algoritmo Apriori para descubrir reglas de asociación e identificar situaciones excepcionales. El primer experimento utilizó datos de muertes infantiles en la región metropolitana de Curitiba-PR. En el segundo experimento, se utilizaron los registros médicos de los pacientes atendidos por enfermeros en la clasificación de riesgo en un hospital privado de Curitiba-PR. El primer estudio identificó 374 reglas generales y el segundo 108 reglas generales, ambas con sus respectivas reglas de excepción. **Conclusión:** la aplicación del Knowledge Discovery in Databases se puede demostrar e implementar en dos escenarios diferentes, con el fin de contribuir a la toma de decisiones por parte del gerente. Se espera que el informe refuerce la importancia de la enseñanza de la informática en enfermería como herramienta de apoyo a la toma de decisiones.

**Palabras clave:** Minería de Datos; Bases de Datos Bibliográficas; Informática Aplicada a la Enfermería; Educación en Enfermería; Toma de Decisiones.

#### Como citar este artigo:

Sartorelli AP, Lopata C, Gomes DC, Cubas MR, Carvalho DR. Knowledge Discovery in Databases para apoio à decisão em atividades vinculadas à Enfermagem. REME - Rev Min Enferm. 2021[citado em \_\_\_\_ \_\_\_\_ \_\_\_\_];25:e-1362. Disponível em: \_\_\_\_\_ DOI: 10.5935/1415.2762.20210010

## INTRODUÇÃO

O armazenamento e a utilização de dados oriundos de registros realizados por profissionais da saúde são desafios que envolvem processos técnico-organizacionais.<sup>1</sup> Entre os profissionais de saúde, a equipe de Enfermagem atua nos cuidados diretos ao paciente e, consequentemente, produz dados assistenciais e gerenciais.

Entretanto, suas habilidades e competências para extrair e utilizar os dados gerados são incipientes. Como gestores, os enfermeiros devem aperfeiçoar sua capacidade para análise de grande quantidade de dados e se apropriar de tecnologias que possibilitem usá-los para apoiar a tomada de decisão.<sup>2</sup>

Entre as alternativas para exploração de dados, as mais frequentes são baseadas em estatística, porém é possível adotar recursos tecnológicos da inteligência artificial para descobrir informações capazes de complementar os resultados estatísticos, em especial a mineração de dados (KDD, do inglês Knowledge Discovery in Databases), que permite a identificação de padrões válidos, novos, úteis e compreensivos.<sup>3</sup> Tais padrões podem ser representados, entre outras formas, por meio de regras de associação. Estudos que relacionam a mineração de dados na área de saúde têm apresentado resultados que podem ser empregados na prevenção de doenças,<sup>4</sup> entretanto, não se encontram relatos de aplicações direta ou indiretamente ligadas à Enfermagem.

Nesse contexto, o presente artigo tem como objetivo relatar experiências de estratégias do uso da mineração de dados em dois cenários de práticas de Enfermagem.

## DESCRIÇÃO DA EXPERIÊNCIA

Em ambas as experiências foi adotado o algoritmo Apriori e sobre o conjunto de regras de associação descoberto foram identificadas as situações de exceção<sup>5</sup> denominadas de descoberta de regras de exceção (DRE). As regras de associação são representadas no formato:  $A \rightarrow C$ , lidas como: "SE [Antecedente], ENTÃO [Consequente]". As regras de exceção são representadas por regras gerais, seguidas das respectivas situações de exceção: SE [A] e [B], ENTÃO NOT [C], lidas como: "SE [A] E [B], ENTÃO [C] NEGADO",<sup>6</sup> sendo que [B] representa um item combinado ao [Antecedente] da regra geral.

Adicionalmente, a cada regra descoberta estão associados dois percentuais (Sup, Conf). O Sup remete à probabilidade de o [Antecedente] ocorrer. A Conf remete à probabilidade condicional de o [Consequente] ocorrer, dada a ocorrência do [Antecedente].

A experiência 1 teve a pesquisa de origem aprovada pelo Comitê de Ética em Pesquisa (CEP) da Pontifícia Universidade Católica do Paraná (PUCPR), sob o Parecer número 2.229.737. A experiência 2 não foi submetida a CEP, por se tratar do uso de base de dados secundária; no entanto, a utilização foi autorizada pela gerência assistencial da instituição, sendo utilizada codificação da identidade dos pacientes e enfermeiros. Antes de iniciar o pré-processamento, a codificação dos dados foi realizada pelo coordenador na unidade. Os nomes dos enfermeiros foram substituídos por códigos de 1 a 12, com a finalidade de garantir anonimato.

### Caso 1

Trata-se de exemplo retirado da dissertação de mestrado intitulada "Mineração de dados para identificação de novas relações entre as variáveis associadas à mortalidade infantil, em município da região metropolitana de Curitiba".<sup>7</sup> Os dados referem-se a uma amostra de 358 registros de óbitos infantis ocorridos de 2010 a 2016, registrados no Sistema de Informação sobre Mortalidade (SIM), investigados e analisados pelo Comitê de Prevenção da Mortalidade Materna e Infantil Municipal.

Foram consideradas 124 variáveis nominais, obtidas a partir de pré-processamento dos dados da ficha-síntese de investigação, a qual se encontra de forma estruturada no SIM, e do roteiro de investigação de óbito infantil e fetal. Os dados oriundos das variáveis foram organizados em planilha Excel. São exemplos das variáveis utilizadas: recomendações e medidas de prevenção; critério de redutibilidade; encaminhamento para serviço de referência; e classificação do recém-nascido. Para a mineração de dados foi utilizada a variável "evitabilidade" do óbito infantil como variável-foco (desfecho), por se tratar de uma variável de governabilidade da gestão pública. Os valores possíveis para a variável evitabilidade são: óbito evitável, inevitável e inconclusivo.

Na mineração de dados foram identificadas 2.592.045 regras de associação. O pós-processamento com DRE resultou em 374 regras gerais para "evitabilidade", com suas respectivas regras de exceção. As regras selecionadas apresentaram significância estatística (teste qui-quadrado).

Na Figura 1 é apresentado um exemplo de regra geral, com suas respectivas regras de exceção. Na regra geral 1 a variável "encaminhada para o sistema de referência" foi associada a "óbito inevitável". Quando essa variável está combinada às variáveis "recomendações e medidas de prevenção" e "critério de redutibilidade", a associação foi modificada para "óbito evitável".

Figura 1 - Regra Geral 1, respectivas regras de exceção e estrutura de leitura da regra

Regra	Antecedente	Consequente	Leitura da regra
Geral 1	SE Encaminhada para o sistema de referência = Sim	ENTÃO Óbito inevitável (29,6%; 52,8%)	29,6% das gestantes foram encaminhadas para o sistema de referência; entre as quais, 52,8% óbitos infantis eram inevitáveis
Exceção 1.1	SE Encaminhada para o sistema de referência = Sim E Recomendações e medidas de prevenção = qualidade da consulta pré-natal	ENTÃO Óbito evitável (3,4%; 100,0%)	3,4% das gestantes foram encaminhadas para o sistema de referência e tiveram recomendações e medidas de prevenção em relação à qualidade da consulta pré-natal; entre as quais, 100,0% dos óbitos infantis eram evitáveis
Exceção 1.2	SE Encaminhada para o sistema de referência = Sim E Critério de redutibilidade = Reduzíveis por adequada atenção à mulher na gestação	ENTÃO Óbito evitável (9,5%; 97,1%)	9,5% das gestantes foram encaminhadas para o sistema de referência e se enquadraram no critério de redutibilidade "reduzível por adequada atenção à mulher na gestação"; entre as quais, 97,1% dos óbitos infantis eram evitáveis

Fonte: dados da pesquisa.

## Caso 2

Foram utilizados dados de relatório gerencial referente a 10.813 prontuários de pacientes atendidos por 12 enfermeiros responsáveis pela classificação de risco em um hospital particular da cidade de Curitiba, Paraná, no mês de março de 2019. O critério para a classificação de risco foi o protocolo Emergency Severity Index,<sup>8</sup> que classifica o risco com base na estimativa do número de recursos necessários para o atendimento, considerando a maior utilização de recursos para os pacientes mais graves.

Para o pré-processamento, os dados oriundos do relatório gerencial foram categorizados. As variáveis utilizadas foram: enfermeiro responsável pela triagem (referência numérica 1 a 12); grupo etário do médico que realizou o atendimento (até 40 anos, 41 a 60 anos ou 61 a 80 anos); desfecho do atendimento (alta, internamento ou óbito); especialidade médica (Clínica Geral, Ortopedia ou Cardiologia); concordância entre a triagem e o

desfecho do atendimento médico (sim ou não); e dia da semana do atendimento (segunda-feira a domingo).

Para a descoberta de regras de associação foi adotada como variável-foco (desfecho) a concordância entre a triagem e o desfecho do atendimento médico.

Foram identificadas 8.099 regras de associação, que resultaram em 108 regras gerais, com suas respectivas regras de exceção.

No Figura 2 é apresentado um exemplo de regra geral, com sua respectiva regra de exceção. Na regra geral o antecedente da regra composto das condições "enfermeiro = 3" e "idade do médico = até 40 anos" foi associado ao consequente "concordância entre a classificação de risco e o desfecho do atendimento médico = sim". Quando, porém, o antecedente da regra geral está complementado com a condição "especialidade médica", o consequente se altera para "concordância entre a classificação de risco e o desfecho do atendimento médico = não".

Figura - 2 - Regra geral 2, respectiva regra de exceção e estrutura de leitura da regra

Regra	Antecedente	Consequente	Leitura da regra
Geral 2	SE Enfermeiro = 3 E Idade do médico = Até 40 anos	ENTÃO Alta médica e concordância entre a classificação de risco e o desfecho do atendimento médico = sim (13,3%; 63,7%)	13,3% dos pacientes foram avaliados pelo enfermeiro 3 e atendidos por médico com idade de até 40 anos; em 63,7% desses casos houve concordância entre a classificação de risco e o desfecho do atendimento médico
Regra de exceção 2.1	SE Especialidade médica = Ortopedia E Enfermeiro = 3 E Idade do médico = Até 40 anos	ENTÃO Alta médica e concordância entre a classificação de risco e o desfecho do atendimento médico = não (3,9%; 52,9%)	3,9% dos pacientes da Ortopedia foram avaliados pelo enfermeiro 3 e atendidos por médico com idade de até 40 anos; em 52,9% desses casos não houve concordância entre a classificação de risco e o desfecho do atendimento médico

Fonte: dados da pesquisa.

## DISCUSSÃO

No experimento 1 o uso do KDD possibilitou a compreensão de variáveis associadas à mortalidade infantil condizentes com a realidade local de um município de médio porte. Embora a mortalidade infantil tenha causas e fatores de risco evidenciados em literatura, a descoberta de regras de exceção utilizando base local permite o planejamento municipal com tomada de decisão assertiva, a partir de intervenções direcionadas e com foco na integralidade das ações.

No exemplo destacado, as regras foram relacionadas tanto aos serviços de referência quanto ao pré-natal. Ao realçar situações em que os óbitos infantis poderiam ter sido evitados, as regras de exceção indicam a necessidade de melhorar o acompanhamento das ações realizadas nos serviços de referência, o que pode servir como alerta aos gestores dos serviços municipais para o direcionamento de recursos em capacitações de equipe especializada. Ainda, pode contribuir para a reflexão de gestores e profissionais de saúde para proporcionar melhorias nas práticas assistenciais.<sup>9</sup>

Por outro lado, as regras indicaram que as medidas de prevenção referentes à qualidade da consulta pré-natal e à adequada atenção à mulher na gestação também podem reduzir os óbitos infantis. Isso reforça os achados do estudo publicado em 2018, no qual 51,3% dos óbitos infantis eram reduzíveis por atenção adequada à mulher na gestação.<sup>10</sup>

Nesse exemplo, a aplicação do processo de KDD poderá reforçar a importância da consulta de Enfermagem no pré-natal de risco habitual e ao mesmo tempo assegurar o princípio da integralidade das ações. Mesmo com o fortalecimento de políticas de universalidade, integralidade e equidade, as ações prioritárias para a redução da mortalidade infantil ainda se encontram insuficientes.<sup>9</sup>

Destaca-se que os profissionais comprometidos na análise das questões referentes à redução da mortalidade infantil não se reduzem aos enfermeiros. Entretanto, a equipe coordenada por eles corresponde a expressivo número de trabalhadores nos serviços da atenção primária à saúde, departamentos de promoção e vigilância à saúde e serviços de referência à atenção à mulher, gestante e recém-nascido. Assim, ao se apropriar de ferramentas como o KDD, o enfermeiro gestor de cuidado ou de serviço potencializa a qualidade do cuidado, ao intervir em variáveis sob sua governabilidade técnico-científica.

No experimento 2, o uso do KDD mostrou potencial para descoberta de conhecimento proveniente de dados relacionados à classificação de risco realizada pelo enfer-

meiro em serviço de emergência. A concordância entre a classificação de risco e o desfecho do atendimento médico pode sugerir um alinhamento de condutas institucionais entre os profissionais, bem como melhor direcionamento de recursos escassos em especialidades críticas e adoção os protocolos preconizados pela instituição. A idade do médico, também presente nas regras de exceção, pode estar relacionada ao maior tempo de experiência.

Estudo que avaliou a aplicação de protocolo de classificação de risco entre enfermeiros demonstrou total concordância nos níveis extremos de necessidade de atendimento – emergência e não urgência –, entretanto, ainda persistem dúvidas sobre a efetividade da ação profissional.<sup>11</sup> No exemplo relatado, os resultados do processo de KDD reforçam situações positivas no que se refere ao atendimento de emergência iniciado pelo enfermeiro, o que poderá justificar a presença de tal profissional.

Por sua vez, as divergências na classificação de risco, obtidas no exemplo, suscitam a discussão de que a atribuição de um grau de risco ao paciente compreende um complexo processo de tomada de decisão, em que é necessária a identificação dos dados a serem coletados e dos instrumentos a serem utilizados, evitando que a tomada de decisão seja conduzida apenas pela experiência e avaliação subjetiva de cada enfermeiro.<sup>12</sup>

Tendo como base os dois exemplos relatados, pode-se inferir que a contribuição da adoção da mineração de dados para apoiar o processo decisório consiste na complementação dos resultados obtidos a partir de testes estatísticos, que no caso dos dois experimentos apresentados poderia ser o teste do qui-quadrado, o qual permite identificar as relações entre variáveis nominais, que representamos nos resultados como “regra geral”.

A mineração de dados torna-se um diferencial a partir do momento em que permite identificar as exceções que, para gestores locais, são mais relevantes do que a associação ou o estabelecimento de relação a partir de testes estatísticos já consagrados em literatura. Ao usar base de dados locais, o resultado favorece a compreensão das relações, em especial quando se discute a leitura da regra e se identificam os nós críticos dos problemas apresentados.

Destaca-se que a estatística e a mineração de dados são complementares. Os exemplos apresentados revelam que o processo KDD explicitou relações que seriam redundantes sem o uso de regras de exceção.

Os relatos apresentados podem se estender a outras situações vivenciadas no cotidiano dos enfermeiros e contribuir para o ensino da informática em Enfermagem nos cursos de graduação e pós-graduação como ferramenta de apoio à decisão.

## CONCLUSÃO

O uso do KDD no cenário relativo à mortalidade infantil auxiliou na compreensão das variáveis associadas e condizentes com a realidade local, podendo ser base para o planejamento e para a tomada de decisão assertiva, com integralidade da atenção à saúde. Sua aplicação reforça a identificação de causas conhecidas da mortalidade infantil, porém explícita em detalhes a relação com variáveis locais.

No caso da classificação de risco pelo atendimento do enfermeiro, o KDD pode sugerir um alinhamento de condutas institucionais entre médicos e enfermeiros, bem como melhor direcionamento de recursos em áreas críticas.

A aplicação do KDD pode ser efetivada em diferentes cenários, contribuindo para a tomada de decisão pelo gestor. E no caso da Enfermagem espera-se que o relato reforce a importância do ensino da Informática em Enfermagem nos cursos de graduação e pós-graduação como ferramenta de apoio à decisão.

A limitação deste relato está na impossibilidade de generalização dos exemplos, em virtude da pequena amostra de registros utilizada no caso 1 e do uso de bases locais.

## REFERÊNCIAS

1. Geilsa SCV, Augusta MD. A gestão da informação em atenção básica de saúde e a qualidade dos registros de Enfermagem. *Investig Enferm*. 2018[citado em 2020 abr. 09];20(2):01-8. Disponível em: <https://pesquisa.bvsalud.org/portal/resource/pt/biblio-995111>
2. Topaz M, Pruinelli L. Big Data and Nursing: implications for the future. *Stud Health Technol Inform* 2017[citado em 2020 jun. 18];232:165-71. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/28106594>
3. Fayyad U, Piatetsky-Shapiro G, Smyth P, Uthurusamy R. *Advances in knowledge discovery and data mining*. American Association for Artificial Intelligence. Menlo Park: MIT Press; 1996.
4. Li Q, Zhang Y, Kang H, Xin Y, Shi C. Mining association rules between stroke risk factors based on the Apriori algorithm. *Technol Health Care*. 2017[citado em 2020 abr. 09];25(S1):197-205. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/28582907>
5. Agrawal R, Srikant R. Fast algorithms for mining association rules. In: *Proceedings of the 20th VLDB Conference*. Santiago, Chile; 1994. p.487-99.
6. Hussain F, Liu H, Suzuki E, Lu H. Exception Rule Mining with a Relative Interestingness Measure. *Proceedings of the 4th Pacific-Asia Conference on Knowledge Discovery and Data Mining PAKDD*, Kyoto, Japão; 2000 Apr. 18-20.
7. Sartorelli AP. Mineração de dados para identificação de novas relações entre variáveis associadas à mortalidade infantil, em um município da região metropolitana de Curitiba [dissertação]. Curitiba: Pontifícia Universidade Católica do Paraná; 2018.
8. Emergency Severity Index (ESI). A triage tool for emergency. Department care. Version 4, Implementation Handbook. Rockville: AHRQ Publication; 2012[citado em 2017 abr. 16]. Disponível em: <https://www.ahrq.gov/sites/default/files/wysiwyg/professionals/systems/hospital/esi/esihandbk.pdf>
9. Pinheiro JMF, Tinoco LS, Rocha ASS, Rodrigues MP, Lyra CO, Ferreira MAF. Atenção à criança no período neonatal: avaliação do pacto de redução da mortalidade neonatal no Rio Grande do Norte, Brasil. *Ciênc Saúde Colet*. 2016[citado em 2020 dez. 27];21(1):243-52. Disponível em: <https://doi.org/10.1590/1413-81232015211.09912014>
10. Oliveira CM, Bonfim CV, Guimarães MJB, Frias PG, Medeiros ZM. Mortalidade infantil: tendência temporal e contribuição da vigilância do óbito. *Acta Paul Enferm*. 2016[citado em 2020 dez. 27];29(3): 282-90. Disponível em: <https://doi.org/10.1590/1982-0194201600040>
11. Malfussi LBH, Bertoncello KCG, Nascimento ERP, Silva SG, Hermida PMV, Jung W. Concordância de um protocolo institucional de avaliação com classificação de risco. *Texto Contexto Enferm*. 2018[citado em 2020 abr. 30];27(1):e4200016. Disponível em: <https://www.scielo.br/pdf/tce/v27n1/0104-0707-tce-27-01-e4200016.pdf>
12. Souza CC. A tomada de decisão do enfermeiro na classificação de risco: qual o embasamento teórico para a prática neste cenário de cuidado? *REME - Rev Min Enferm*. 2015[citado em 2020 abr. 08];19(2):09-11. Disponível em: <http://www.dx.doi.org/10.5935/1415-2762.20150021>