Caracterização da Mortalidade Infantil dos Estados de Santa Catarina e Amapá Utilizando Mineração de Dados

Wanderson L. G. Soares, Patrícia Lima, Luis E. Zárate, Mark A. Junho Song e Cristiane N. Nobre

Pontifícia Universidade Católica de Minas Gerais
wanderson.lg.soares@gmail.com, patricia.lima.909576@sga.pucminas.br, {song, nobre, zarate}@pucminas.br

Abstract. The objective of this article is to use the concepts of knowledge discovery in databases, specifically the concepts of machine learning in the data mining phase, to characterize infant mortality in the state of Santa Catarina (with lower infant mortality rate) and in the state of Amapá (with the highest infant mortality rate). In this way, the classifiers J48, JRip and Random Forest were used and a brief comparison was made between the results obtained by the classifiers in both states. In addition, the database was preprocessed, which includes attribute selection and balancing, the application of data mining techniques and the analysis of the results of the respective models.

Categories and Subject Descriptors: H.2.8 [**Database Management**]: Database Applications; I.2.6 [**Artificial Intelligence**]: Learning

Keywords: Classification, Data Mining, Infant Mortality, Machine Learning.

1. INTRODUÇÃO

O estudo do índice de mortalidade infantil pode revelar detalhes sobre aspectos que precisam ser aprimorados na população e trata-se de um fator decisivo sobre o desenvolvimento do estado. A Mortalidade Infantil apresenta, sob o aspecto científico e social, uma forma de avaliar tanto a questão comunitária, quanto as medidas de saúde adotadas em uma determinada região e trata-se de um evento que aflige o mundo inteiro [Black et al. 2010].

No trabalho de Hernandez et al. (2011) foi relatado que a *MI* possui aspectos associados aos problemas de desigualdade social. Diante deste contexto, a taxa de *MI* é um parâmetro relevante que poderá revelar as condições de saúde de uma determinada população [Vianna et al. 2010], e também com o acesso dessa população aos serviços de saúde prestados [Brasil 2009].

Visando obter informações sobre *MI*, o governo brasileiro implantou o Sistema de Informações sobre Mortalidade (*SIM*) em 1975 e o Sistema de Informações sobre Nascidos Vivos (*SINASC*) em 1990. O *SIM* é uma base de dados que inclui todos os registros sobre mortalidade e o *SINASC* reúne informações sobre os nascimentos em todo o território brasileiro.

Com relação às bases de dados do *SIM* e do *SINASC*, o processo de Descoberta de Conhecimento em Bases de Dados *(Knowledge Discovery in Databases - KDD)* é uma abordagem que possibilita a inferência de conhecimento a partir de uma grande base de dados [Felix 1998]. Segundo VIANNA et al. (2010), a utilização das técnicas de *KDD* é bem satisfatória na obtenção de conhecimento e relata-se que as técnicas de aprendizado de máquina estão entre as mais utilizadas no processo de mineração de dados.

O objetivo deste trabalho é caracterizar a *MI* nos estados de Santa Catarina (com a menor taxa de

MI) e do Amapá (que apresenta a maior taxa de *MI*) com a utilização dos 03 (três) classificadores: *J48* com o plugin *VTJ48*, *JRip* e *Random Forest*, disponíveis no software *Waikato Environment for Knowledge Analysis (WEKA)*. A justificativa pela escolha desses três classificadores está no fato de que o *J48* e o *JRip* descrevem as regras de classificação, enquanto o *Random Forest* categoriza os atributos mais relevantes. Assim, este trabalho objetiva avaliar as regras e os atributos indicados para se identificar os principais fatores que contribuem para a mortalidade infantil nos dois estados considerados.

Este artigo está organizado da seguinte maneira: A Seção 2 contém o referencial teórico englobando informações sobre mortalidade infantil, mineração de dados e os respectivos classificadores: *J48* com o plugin *VTJ48*, *JRip* e *Random Forest*. A Seção 3 apresenta os trabalhos relacionados. A Seção 4 descreve materiais e métodos, descrevendo a base de dados, o pré-processamento e as métricas de avaliação. A Seção 5 apresenta os resultados e discussões a partir das métricas de avaliação. A Seção 6 apresenta as considerações finais e conclui com os trabalhos futuros.

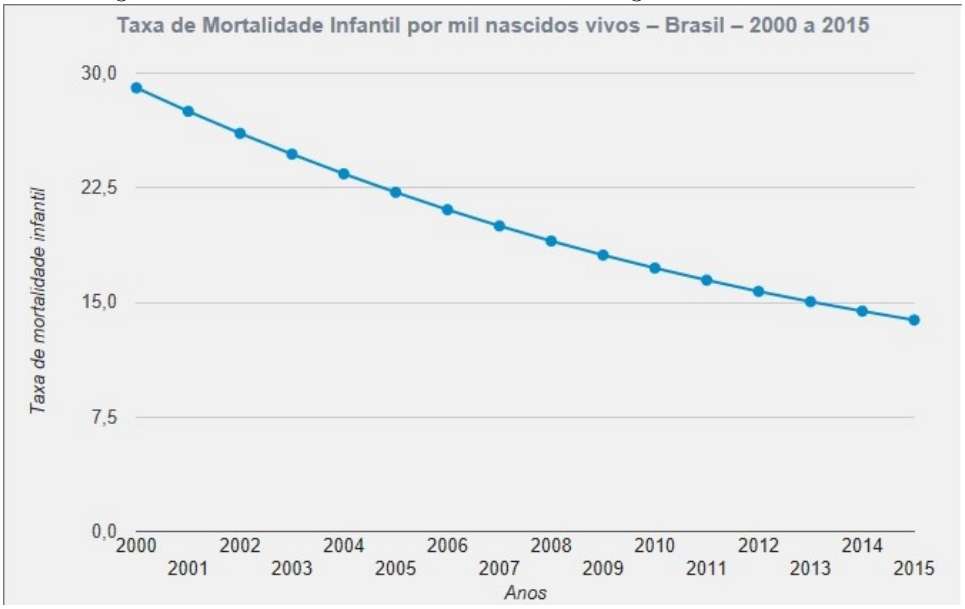
2. REFERENCIAL TEÓRICO

A fundamentação teórica deste trabalho consiste em abordar os conceitos sobre a *MI* e o processo de *KDD*.

2.1 Mortalidade Infantil

A *MI* engloba os óbitos dos seguintes períodos: neonatal precoce (0-6 dias de vida), neonatal tardio (7-27 dias de vida) e pós-neonatal (28 e 364 dias de vida). Constata-se que durante o período neonatal precoce ocorrem por volta de 50% dos óbitos infantis, podendo chegar até por volta de 66% durante o período neonatal tardio [UNICEF et al. 2008].

Fig. 1. Taxa de Mortalidade Infantil no Brasil ao longo dos anos 2000 a 2015



Fonte: IBGE, Projeção da População do Brasil - 2013.

Segundo o Instituto Brasileiro de Geografia e Estatística (IBGE), a taxa de *MI* no Brasil tem apresentado um declínio significativo e tem se mostrado associada com os fatores sociais e econômicos

do nosso país [Goldani et al. 2001]. Segundo Costa et al. (2003), a redução da *MI* observada nos anos de 1980 esteve relacionada ao declínio da fecundidade. Relata-se que a taxa de *MI* foi de 52,02/1.000 em 1989, para 35,57/1.000 em 1998. Na Figura 1, releva-se que a taxa de *MI* foi de 29,02/1.000 no ano 2000, para 13,8/1.000 em 2015.

2.2 Mineração de Dados e Descoberta de Conhecimento

A mineração de dados pode ser descrita como a atividade de extrair conhecimento e/ou padrões a partir de uma grande quantidade de dados [Quilici-Gonzalez 2015]. Grande parte da literatura trata a mineração de dados como descoberta de conhecimento em bases de dados (*KDD – Knowledge Discorey in Databases*) e tem autores que consideram a mineração de dados como sendo uma etapa do processo do *KDD* [Carvalho et al. 2011]. Fayyad et al. (1996) relatam que as etapas de um processo *KDD* são as seguintes: Seleção, Limpeza e integração de dados, Transformação dos dados, Mineração de Dados, Avaliação e apresentação dos resultados.

2.2.1 *Técnicas de Balanceamento.* No mundo real, a quantidade de instâncias de diferentes classes poderá variar [Prati et al. 2003], por exemplo: poderá existir uma classe A com uma quantidade de instâncias muito superior quando comparada com a classe B, sendo assim a classe A será majoritária neste exemplo. É importante salientar que o desbalanceamento pode afetar negativamente o resultado de algoritmo baseado em aprendizado de máquina [Carvalho et al. 2011]. Diante de dados desbalanceados é importante realizar o balanceamento a partir das seguintes abordagens:

—*Oversampling*: consiste na replicação de instâncias da classe minoritária visando realizar o balanceamento, mas o acréscimo de instâncias poderá incorporar situações que nunca ocorreram na prática

—*Undersampling*: consiste na eliminação de instâncias da classe majoritária visando realizar o balanceamento, mas isso poderá levar à eliminação de dados relevantes que poderão comprometer a indução do modelo

2.3 Classificadores utilizados

Este trabalho utiliza os seguintes classificadores: *J48*, *JRip* e *Random Forest*.

2.3.1 *J48 com o plugin VTJ48.* Este trabalho utiliza o algoritmo *J48*, implementação em Java do algoritmo C4.5 (QUINLAN, 1993) na plataforma *WEKA*. Para ajustar os parâmetros do *J48*, foi utilizado o pluging *VTJ48*[1]. O algoritmo C4.5 visa a geração de árvores de decisão permitindo o tratamento de atributos numéricos e/ou nominais. Durante o treinamento, a cada nó o algoritmo seleciona um atributo que melhor subdivide o conjunto das amostras [Quinlan 2014].

2.3.2 *JRip.* O algoritmo *Repeated Incremental Pruning to Produce Error Reduction* (*RIPPER*) refere-se à versão otimizada do algoritmo *Incremental Reduced Error Pruning* (*IREP*) [Cohen 1995]. O algoritmo (*RIPPER*) adota a abordagem de poda reduzida visando a redução de erros e a geração de regras adequadas. Este trabalho utiliza o algoritmo *JRip*, implementação em Java do algoritmo (*RIPPER*).

2.3.3 *Random Forest.* O algoritmo consiste em um grande número de árvores e adota a abordagem do voto majoritário para fazer a classificação. O respectivo classificador apresenta resultado satisfatório mesmo com a presença de ruído/outliers [Khoshgoftaar et al. 2007] e as árvores do *Random Forest* são caracterizadas por terem suas entradas definidas de uma forma aleatória [Breiman 2001].

[1]Plugin disponível em: http://www.ri.fzv.um.si/vtj48/. Acessado em: 09 jun. 2018.

3. TRABALHOS RELACIONADOS

Todos os trabalhos descritos nesta seção estão envolvidos com a aplicação de técnicas de *KDD* na área da saúde, os quais incluem o estudo sobre a *MI* e a tentativa de descoberta de padrões que possam resultar em alguma medida de intervenção para diminuir as taxas de *MI*.

Para Oliveira et al [2001], a mineração de dados pode ser utilizada como ferramenta para desenvolver um modelo de prevenção da mortalidade infantil. Os autores utilizaram técnicas de classificação associadas ao processo de *KDD* para traçar o perfil de recém-nascidos e identificar quais variáveis estão associadas à sua mortalidade. Os resultados estatísticos apresentam uma forte correlação ao peso do bebê, ao nível de apgar (exame que avalia o nível de adaptação do bebê à vida fora do útero) do primeiro e quinto minuto de vida, à duração da gestação (em semanas) e ao tipo de gravidez (única, dupla, etc).

Buscando trabalhos que corroborassem com a aplicação das técnicas de *KDD*, [Barcellos et al. 2002] analisou a situação atual do geoprocessamento e da análise de dados na rede de saúde pública brasileira. Os autores concluíram que as técnicas computacionais são de grande auxílio, devem ser cada vez mais utilizadas como medidas de análises na área da saúde e que são necessários investimentos para capacitação de pessoas para realização deste trabalho.

No trabalho de Kitsantas et al [2006] o objetivo era identificar os subgrupos de mulheres com alto risco de desenvolver uma gestação onde o bebê nasça com baixo peso. Os dados são de sete regiões geográficas da Flórida, em que aplicando técnicas de mineração de dados foi possível identificar vários subgrupos de alto risco, inclusive o baixo peso ao nascer que era a hipótese principal.

Segundo Vianna et al. [2010], são identificados padrões de características materno-fetais na predição da *MI* utilizando mineração de dados. Para o estudo foi realizada a integração das bases de dados do SINASC, do SIM e do SIMI (Sistema de Investigação da Mortalidade Infantil do Paraná) com relação ao período de 2000 a 2004, a fim de reunir, através da aplicação dessas técnicas de mineração, um conjunto de ações voltadas às regras mais relacionadas à *MI*. Desta forma, este artigo concluiu que devem ocorrer ações voltadas para mães adolescentes (principalmente as que já têm outro filho), mães com problemas na gestação, mães com filhos que possuem baixo peso ao nascer e com pós-datismo.

4. MATERIAIS E MÉTODOS

Para realização da mineração de dados, a fim de caracterizar a *MI* nos estados de Santa Catarina (SC) e Amapá (AP) no ano de 2015, foram seguidas as subsequentes etapas: escolha da base de dados, pré-processamento da base de dados, utilização dos três classificadores: *J48* com o plugin *VTJ48*, *JRip* e *Random Forest*, análise dos resultados das árvores de cada estado e comparação de resultados.

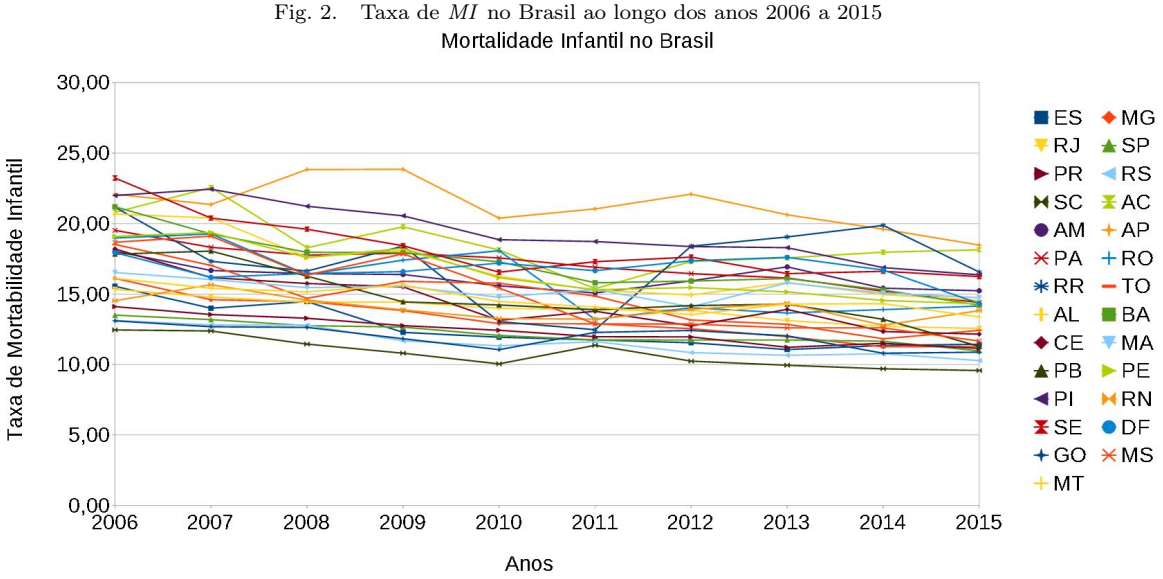
4.1 Base de Dados

A base de dados sobre mortalidade infantil foi obtida no site do DATASUS considerando os dados do período de 2006 até 2015 e englobou neste trabalho os dados do SINASC e do SIM. Para cada ano e para cada estado, o DATASUS disponibiliza um arquivo com os dados do SIM e outro arquivo com os dados SINASC.

A taxa de *MI* é apresentada na Figura 2 para todos os estados brasileiros, inclusive o Distrito Federal, considerando o período de 2006 até 2015. A taxa de *MI* é o número de óbitos de menores de um ano de idade, por mil nascidos vivos, por cada estado e ano considerado. É possível visualizar que os estados de Santa Catarina e do Amapá apresentam, respectivamente, as taxas de *MI* menor e maior, quando comparados com os demais estados do Brasil, incluindo o Distrito Federal (Figura 2).

No pré-processamento das bases foi realizada a integração dos dados do *SIM* e do *SINASC*, mantendo os atributos comuns e unificando o atributo Local de Ocorrência do Óbito (LOCOCOROBI -

Fig. 2. Taxa de *MI* no Brasil ao longo dos anos 2006 a 2015



SIM) com o atributo Local Ocorrência do Nascimento (LOCOCORNASC - *SINASC*), dando origem ao campo Local de Ocorrência do Nascimento ou Óbito na base de dados unificada. Como resultado obteve-se os atributos listados na Tabela I.

Tabela I. Campos e Descrições dos Atributos Mantidos Após o Pré-Processamento

Atributos	**Descrição**
Idade da Mãe	Em anos
Escolaridade da Mãe	Em anos
Quantidade de filhos vivos	Numérico contínuo
Quantidade de filhos mortos	Numérico contínuo
Gravidez	Única, dupla, tripla ou mais
Gestação	Em semanas
Parto	Ignorado, normal, cesáreo
Peso	Em gramas ao nascer
Local de ocorrência do nascimento ou óbito	Ignorado, hospital, outro estabelecimento de saúde, domicílio, via pública, outros
Sexo	Ignorado, masculino, feminino
Raça	Branca, Preta, amarela, parda, indígena
Classificação	Vivo, óbito infantil

Todos os registros que continham algum atributo com valor ausente e também um conjunto de registros relacionados às inconsistências constatadas, ambos conjuntos de registros foram retirados da base para criação do modelo, a fim de não afetarem os resultados da árvore de decisão. Foi realizado este pré-processamento com o objetivo de gerar uma única base de dados que englobasse as classes Vivo e óbito infantil. Após unificadas, as bases continham instâncias desproporcionais quanto às classes (Tabela II), sendo necessária a aplicação de técnicas de balanceamento de dados.

Tabela II. Dimensões das Bases de Dados Unificadas, antes e após o balanceamento

	Antes do balanceamento		Após o balanceamento	
	Santa Catarina	Amapá	Santa Catarina	Amapá
Vivo	74394	11885	331	87
Óbito Infantil	331	87	331	87

4.2 Pré-processamento dos Dados

Foi constatado um desbalanceamento entre a classe Vivo (classe majoritária) e a classe *Óbito Infantil* (classe minoritária). É importante salientar que o desbalanceamento pode afetar negativamente o resultado de algoritmos baseados em aprendizado de máquina. Diante deste fato, foi utilizado o balanceamento de classes utilizando-se a abordagem *undersampling*. Este balanceamento é caracterizado pelo fator de eliminar dados da classe majoritária produzindo um subconjunto aleatório de dados visando que a classe majoritária passe a ficar com o mesmo número de instâncias da classe minoritária.

4.3 Métricas de Avaliação

Para avaliação da qualidade dos modelos obtidos, foram utilizadas as métricas de precisão, sensibilidade e F-measure.

Precisão (Equação 1) mede a proporção de instâncias classificadas em determinada classe que são realmente da classe.

$$Pr = \frac{VP}{VP + FP} \tag{1}$$

Sensibilidade (Equação 2) mede a proporção de instâncias corretamente classificadas, dentre todas as instâncias de uma classe.

$$Sen = \frac{VP}{VP + FN} \tag{2}$$

F-measure (Equação 3) representa a média harmônica entre precisão e sensibilidade.

$$F - measure = \frac{(w+1) * Sen * Pr}{Sen + w * Pr} \tag{3}$$

onde VP= Verdadeiros Positivos, FP=Falsos Positivos e FN = Falsos Negativos.

Para definir os conjuntos de treinamento e teste, foi utilizado o método *cross-validation* de 10 dobras. Este método tem como objetivo avaliar a capacidade de generalização de um modelo [Kohavi et al. 1995].

5. RESULTADOS E DISCUSSÕES

As Tabelas III, IV e V apresentam as métricas de avaliação da qualidade dos modelos gerados, respectivamente, pelos classificadores *J48* com o plugin *VTJ48*, *JRip* e *Random Forest*, assim como as análises dos resultados.

Tabela III. Métricas de avaliação e qualidade do modelo: *J48* com o plugin *VTJ48*

	Santa Catarina		**Amapá**	
	Vivo	Óbito Infantil	Vivo	Óbito Infantil
Precisão	0,889	0,936	0,812	0,932
Sensibilidade	0,940	0,882	0,943	0,782
F-Measure	0,913	0,908	0,872	0,850

A partir das árvores de decisão geradas pelo classificador *J48* com o plugin *VTJ48* para os estados, foi constatado que o atributo peso se mostra como o mais relevante em ambos os casos, classificando

Tabela IV. Métricas de avaliação e qualidade do modelo: *JRip*

	Santa Catarina		**Amapá**	
	Vivo	Óbito Infantil	Vivo	Óbito Infantil
Precisão	0,871	0,914	0,787	0,970
Sensibilidade	0,918	0,864	0,977	0,736
F-Measure	0,894	0,888	0,872	0,837

Tabela V. Métricas de avaliação e qualidade do modelo: *Random Forest*

	Santa Catarina		**Amapá**	
	Vivo	Óbito Infantil	Vivo	Óbito Infantil
Precisão	0,898	0,925	0,835	0,922
Sensibilidade	0,927	0,894	0,931	0,816
F-Measure	0,912	0,909	0,880	0,866

260 instâncias de Santa Catarina e 65 do Amapá. No estado de SC, os bebês que nascem abaixo de 2.215 gramas são muito propensos ao óbito, no estado do AP esse número sobe para 2.469 gramas.

Com relação ao classificador *JRip*, as principais regras geradas por este classificador consideraram o atributo peso como sendo o mais relevante, e analisando os resultados pode-se verificar que o peso mínimo para classificar como Vivo nos estados de Santa Catarina e Amapá foram, respectivamente 2.220 e 2.469 gramas. Assim, os resultados do classificador *JRip* estão harmônicos com os resultados do classificador *J48*.

Em se tratando do *Random Forest*, o respectivo classificador considerou para o estado de Santa Catarina os seguintes três atributos mais relevantes dentro da seguinte ordem: idade da mãe, peso, quantidade de filhos vivos. Já para o estado do Amapá, os atributos mais relevantes foram: idade da mãe, peso, quantidade de filhos mortos. Com base nos resultados do classificador *Random Forest*, é possível considerar que os atributos comuns mais relevantes, tanto para SC quanto para AP, foram: a idade da mãe e o peso. Considerando o número de nodos que utilizam esses atributos, os que se destacam são o peso e a idade da mãe.

Com o intuito de avaliar o modelo utilizado após o balanceamento da base de dados, foi analisada a qualidade a partir das instâncias resultantes do pré-processamento sem balanceamento que não participaram do *cross-validation* e os resultados com os percentuais de instâncias classificadas corretamente do *J48* foram de 90% para SC e 82% para o AP. Já o *JRip* obteve 88% para SC e 91% para o AP. Finalmente, o *Random Forest* apresentou 92% para SC e 89% para o AP. Esses resultados demonstram que a maioria das instâncias foram corretamente classificadas e, sendo assim, os modelos gerados pelos classificadores: *J48* com o plugin *VTJ48*, *JRip* e *Random Forest* são satisfatórios. É importante ressaltar que o atributo peso foi considerado relevante analisando os resultados dos três classificadores: *J48*, *JRip* e *Random Forest*. Assim, estes resultados corroboram com o trabalho desenvolvido por Barbas et al. (2009), onde bebês com menos de 2.500 gramas são considerados com alto risco, ou seja, possuem alta probabilidade de óbito antes de completar um ano de vida.

6. CONSIDERAÇÕES FINAIS

Através dos resultados gerados por cada classificador: *J48* com o plugin *VTJ48*, *JRip* e *Random Forest*, pode-se observar que o atributo peso é relevante e que a *MI* é caracterizada basicamente por bebês com peso inferior a 2.215 gramas em SC e 2.469 gramas no AP. Como medidas preventivas, visando minimizar a taxa de *MI*, é necessário que hajam mais investimentos na área da saúde, já que o peso pode estar diretamente relacionado ào acompanhamento pré-natal da gestante.

Como proposta de trabalhos futuros, seria interessante aplicar as técnicas de *KDD* em todos os estados do território brasileiro e em todos os anos disponibilizados pelo *DATASUS* visando caracterizar

a *MI* em todos os demais estados brasileiros e nortear os investimentos públicos para a diminuição da taxa de *MI* no país.

Agradecimentos

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior(CAPES) pelo apoio financeiro.

REFERENCES

Barbas, D. d. S., Costa, A. J. L., Luiz, R. R., and Kale, P. L. Determinantes do peso insuficiente e do baixo peso ao nascer na cidade do rio de janeiro, brasil, 2001. *Epidemiologia e Serviços de Saúde* 18 (2): 161–170, 2009.

Barcellos, C. d. C., Ramalho, W. M., et al. Situação atual do geoprocessamento e da análise de dados espaciais em saúde no brasil, 2002.

Black, R. E., Cousens, S., Johnson, H. L., Lawn, J. E., Rudan, I., Bassani, D. G., Jha, P., Campbell, H., Walker, C. F., Cibulskis, R., et al. Global, regional, and national causes of child mortality in 2008: a systematic analysis. *The lancet* 375 (9730): 1969–1987, 2010.

Brasil, M. d. S. Manual de vigilância do óbito infantil e fetal e do comitê de prevenção do óbito infantil e fetal, 2009.

Breiman, L. Random forests. *Machine learning* 45 (1): 5–32, 2001.

Carvalho, A., Faceli, K., LORENA, A., and GAMA, J. Inteligência artificial–uma abordagem de aprendizado de máquina. *Rio de Janeiro: LTC*, 2011.

Cohen, W. W. Fast effective rule induction. In *Machine Learning Proceedings 1995*. Elsevier, pp. 115–123, 1995.

Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. The kdd process for extracting useful knowledge from volumes of data. *Communications of the ACM* 39 (11): 27–34, 1996.

Felix, L. C. M. *Data mining no processo de extraçao de conhecimento de bases de dados*. Ph.D. thesis, Universidade de São Paulo, 1998.

Goldani, M. Z., Barbieri, M. A., Bettiol, H., Barbieri, M. R., and Tomkins, A. Infant mortality rates according to socioeconomic status in a Brazilian city. *Revista de Saúde Pública* vol. 35, pp. 256 – 261, 06, 2001.

Hernandez, A. R., Silva, C. H. d., Agranonik, M., Quadros, F. M. d., and Goldani, M. Z. Análise de tendências das taxas de mortalidade infantil e de seus fatores de risco na cidade de porto alegre, rio grande do sul, brasil, no período de 1996 a 2008. *Cadernos de Saúde Pública* vol. 27, pp. 2188–2196, 2011.

Khoshgoftaar, T. M., Golawala, M., and Van Hulse, J. An empirical study of learning from imbalanced data using random forest. vol. 2, pp. 310–317, 2007.

Kitsantas, P., Hollander, M., and Li, L. Using classification trees to assess low birth weight outcomes. *Artificial intelligence in medicine* 38 (3): 275–289, 2006.

Kohavi, R. et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*. Vol. 14. Montreal, Canada, pp. 1137–1145, 1995.

Oliveira, I. T. C. d. et al. Aplicação de data mining na busca de um modelo de prevenção da mortalidade infantil, 2001.

Prati, R. C., Batista, G., and Monard, M. C. Uma experiência no balanceamento artificial de conjuntos de dados para aprendizado com classes desbalanceadas utilizando análise roc. In *Proc. of the Workshop on Advances & Trends in AI for Problem Solving*. Vol. 1. pp. 28–33, 2003.

Quilici-Gonzalez, José Artur de Assis Zampirolli, F. *Sistemas inteligentes e mineração de dados*, 2015.

Quinlan, J. R. *C4. 5: programs for machine learning*. Elsevier, 2014.

UNICEF, N. B. et al. Disponível em: http://www. unicef. org/brazil/pt/. *Acesso em: junho de 2018*, 2008.

Vianna, R. C. X. F., Moro, C. M. C. d. B., Moysés, S. J., Carvalho, D., and Nievola, J. C. Mineração de dados e características da mortalidade infantil. *Cadernos de Saúde Pública* vol. 26, pp. 535–542, 2010.