

# SISTEMA DE PREDIÇÃO DE ALARMES EM PROCESSOS INDUSTRIAIS POR CLASSIFICAÇÃO NÃO-SUPERVISIONADA

SERGIO H. BRAUNSTEIN<sup>1</sup>, ANDRE P. LERM<sup>2</sup>, RAFAEL A. R. LERM<sup>3</sup>, ADRIANO V. WERHLI<sup>4</sup>,  
SILVIA S. C. BOTELHO<sup>5</sup>, EDWALDO O. LIPPE<sup>6</sup>

## RESUMO

Um sistema de predição de alarmes com a finalidade de auxiliar a implantação de uma política de manutenção preditiva industrial e de constituir-se em uma ferramenta gerencial de apoio à tomada de decisão é proposto neste trabalho. O sistema adquire leituras de diversos sensores instalados na planta, extrai suas características e avalia a saúde do equipamento. O diagnóstico e prognóstico implica a classificação das condições de operação da planta. Técnicas de árvores de regressão e classificação não-supervisionada são utilizadas neste artigo. Uma amostra das medições de 73 variáveis feitas por sensores instalados em uma usina hidrelétrica foi utilizada para testar e validar a proposta. As medições foram amostradas em um período de 15 meses.

**PALAVRAS-CHAVE:** árvores de regressão, classificação não-supervisionada, manutenção preditiva, séries temporais.

<sup>1</sup>Mestrando do Programa de Pós-Graduação em Modelagem Computacional – FURG; shb@pelotas.ifsul.edu.br

<sup>2</sup>Professor do Instituto Federal Sul-rio-grandense – Pelotas; Doutor em Eng. Elétrica – aaplerm@gmail.com

<sup>3</sup>Acadêmico do Curso de Engenharia de Computação – FURG; rafaellerm@gmail.com

<sup>4</sup>Professor do Pós-Graduação em Modelagem Computacional – FURG; Doutor em Informática – werhli@gmail.com

<sup>5</sup>Professora do Pós-Graduação em Modelagem Computacional – FURG; Doutora em Informática e Telecom. – silviacb@furg.br

<sup>6</sup>Gerente de Engenharia e Planejamento – AES Tietê SA; edwaldo.lippe@aes.com

## ABSTRACT

### INDUSTRIAL PROCESSES ALARM PREDICTION USING NON-SUPERVISED CLASSIFICATION

In this work an alarm prediction system is proposed. Its main aims are to contribute to the establishment of predictive industrial maintenance guidelines and to produce a management decision support tool. The proposed system obtains readings from many sensors that are installed in the industrial plant, extract its characteristics and evaluates the equipment's health. The diagnosis and prognosis implies in a classification of the industrial plant's operational condition. Classification and regression trees are applied in this paper. A measurement sample from 73 sensors installed in a hydroelectric power plant is utilized to test and validate the proposed methodology. The measurements were obtained in a 15 months period.

**KEYWORDS:** atheoretical regression trees, non-supervised classification, predictive maintenance, time series.

## 1 – INTRODUÇÃO

Nos tempos atuais, a busca pelo aumento da competitividade justifica investimentos no processo de produção de forma a alcançar elevados índices de confiabilidade e de desempenho do sistema de produção. A manutenção industrial constitui-se em elemento fundamental para a obtenção de melhores resultados operacionais.

Três principais políticas de manutenção podem ser elencadas [16]:

Manutenção corretiva: manutenção voltada à reparação dos equipamentos após a falha ter ocorrido.

Manutenção preventiva: a inspeção, reparação ou substituição do equipamento é feita em agendamentos pré-determinados antes da ocorrência da falha, trazendo consigo uma série de vantagens, entre elas a minimização de imprevistos e redução de custos de horas extras.

Manutenção preditiva: também conhecida como manutenção baseada em condições (*condition-based maintenance - CBM*), além de a atividade de manutenção ser programada antes da ocorrência de falha da máquina, considera as informações coletadas sobre suas condições. O monitoramento faz uso das informações sensoriais para identificar o estado de uma máquina ou equipamento e estimar seu tempo de vida residual. Referência [7] aborda que os principais elementos da manutenção preditiva são monitoramento/previsão, diagnóstico/prognóstico e tomada de decisão/manutenção.

O presente trabalho é resultado do Programa de Pesquisa e Desenvolvimento (P&D) da empresa AES Tietê juntamente com Instituto Federal Sul rio-grandense e Universidade Federal do Rio Grande, o qual propõe um sistema computacional de supervisão e diagnóstico voltado à predição de diferentes tipos de alarmes aplicáveis em plantas de geração hidroelétrica a partir de um conjunto de variáveis de sinais de entrada monitorados.

As diversas variáveis monitoradas são responsáveis pelos disparos de  $n$  alarmes, sendo que um único alarme pode estar associado a mais de uma variável ou ainda, uma variável pode estar vinculada a mais de um alarme. Mineração de dados deverá ser realizada identificando o subconjunto de sinais de entrada associado a cada alarme. Em função do mapeamento estabelecido e do monitoramento dos sinais adequados, o sistema deverá prever as possíveis falhas (alarmes) no processo de produção em tempo hábil de forma a possibilitar ações de manutenção e gerenciais para evitá-las.

Nesta proposta global de trabalho, o presente artigo se detém nas questões de mineração de dados e de classificação das assinaturas normais e anormais das séries temporais. O trabalho está estruturado em quatro seções, a Seção 2 apresenta trabalhos relacionados, Seção 3 aborda a arquitetura geral do sistema e os métodos adotados para o seu desenvolvimento, seguindo na Seção 4 com resultados e discussão e, por fim, na Seção 5 apresentam-se as conclusões.

## **2 – TRABALHOS RELACIONADOS**

Em se tratando de manutenção preditiva, a especificação *OSA-CBM (Open Systems Architecture for Condition-based Maintenance)*, arquitetura padrão para mover informações em sistemas de manutenção baseada em condições, divide o

sistema em seis blocos seqüenciais, sendo Aquisição de Dados (Módulo dos Sensores), Manipulação de Dados (Processamento de Sinais), Monitoramento das Condições, Avaliação da Saúde, Prognóstico e Apoio à Tomada de Decisão [5], [12].

Técnicas de fusão de múltiplos sensores são utilizadas para combinar todas as informações parciais oriundas de diferentes sensores a fim de se obter uma maior precisão de diagnóstico e prognóstico. Tais técnicas podem ser agrupadas em três principais níveis dependendo do estágio de processamento em que elas ocorrem, sendo Fusão no Nível de Dados, Fusão no Nível de Características e Fusão no Nível de Decisão [9], [6].

Na literatura existe uma enorme gama de técnicas e algoritmos a serem empregados na fusão de sensores, entre elas Média Ponderada, Filtro de Kalman, Dempster-Shafer, Redes Neurais Artificiais e outros [6].

Referência [14] utiliza uma fusão no nível de decisão com a abordagem Dempster-Shafer para diagnóstico de falta em motor a diesel de acordo com [9].

Referência [4] aplica a média ponderada dos valores de confiança individuais dos sensores para obtenção do valor de confiança da fusão de sensores em um experimento realizado em uma máquina de solda.

Diversas ferramentas são utilizadas para o desenvolvimento da especificação OSA-CBM. Uma síntese destas é apresentada na Figura 1.

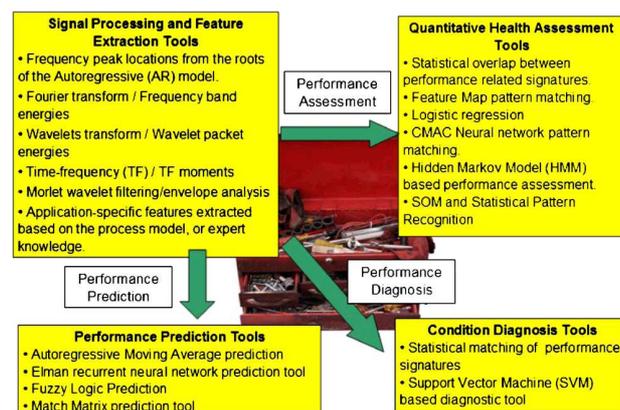


FIGURA 1 – Ferramentas para Processamento de Sinais, Avaliação de Saúde, Diagnóstico e Prognóstico [10]

Referência [11] utiliza k-means para apresentar uma solução alternativa para o problema da continuidade do serviço em sistemas de distribuição de energia associada à localização de falha.

### 3 – METODOLOGIA

Técnicas estatísticas podem ser utilizadas para redução dimensional, reduzindo o volume de informações a serem introduzidas nos classificadores [13]. Métodos de árvores são particularmente adequados para tarefas de mineração de dados, onde muitas vezes há pouco conhecimento a priori do sistema, bem como poucas informações sobre o relacionamento das variáveis existentes [15].

A proposta deste trabalho faz uso de métodos de árvores de regressão no auxílio ao mapeamento dos sinais monitorados nos alarmes a serem previstos. A arquitetura do sistema é composta pelos módulos Extração de Médias, Fornecimento de Utilidades e Previsão de Alarmes.

#### 3.1 – Extração de Médias

Este módulo é responsável pelas quebras estruturais nas séries temporais das variáveis monitoradas, tendo como objetivo identificar os subconjuntos presentes nas séries. O problema exige conhecer como se comportam as variáveis, analisar as alterações nos níveis do seu sinal antes da ocorrência dos  $n$  disparos de um alarme, a fim de tentar extrair algum comportamento padrão destas antes do alarme ocorrer.

No intuito de focar este problema será utilizado o método de árvores de regressão de mínimos quadrados denominado *Atheoretical Regression Trees* (ART). O método ART recorre a um uso seqüencial do método de Fisher de exata otimização. O algoritmo de Fisher é projetado para situações em que os dados estão ordenados e os grupos consistem em intervalos de dados. Referência [3] aborda que na árvore de regressão de mínimos quadrados um nodo  $t$  é dividido em esquerda e direita descendentes, ( $t_e$  e  $t_d$ ), para reduzir o desvio da variável resposta. Assim, o algoritmo seleciona a divisão  $s$  para a qual

$$SS(t) - [SS(t_e) + SS(t_d)] \quad (1)$$

é máximo, onde

$$SS(t) = \sum_{y_i \in t} (y_i - \bar{y}(t))^2, (i = 1, \dots, n) \quad (2)$$

é a soma dos quadrados para o nodo  $t$ ,  $SS(t_e)$  e  $SS(t_d)$  são a soma dos quadrados da esquerda e direita descendentes, respectivamente. Tendo em vista que  $t_e$  e  $t_d$  são sucessivas partições de  $t$  e  $SS(t)$  eles podem ser considerados como sendo a soma dos quadrados do nodo  $t$ . Assim, o critério de divisão consiste em minimizar, em todas as partições binárias de  $t$ , a soma dos quadrados do grupo. Então,

$$MSS_{y|s}(t) = [SS(t_e) + SS(t_d)] \quad (3)$$

Uma vez que a partição binária de um nodo é encontrada, o processo de divisão é aplicado separadamente e recursivamente para cada subconjunto até que estes cheguem a uma dimensão mínima ou nenhuma melhora do critério possa ser alcançada.

Uma janela temporal define o conjunto de dados que será utilizado para a análise e esta deve ser escolhida em função do menor intervalo de tempo observado para todos os disparos do alarme. A janela corresponde ao menor intervalo de tempo entre a saída de um alarme (correção da causa geradora) e o novo disparo deste.

### 3.2– Fornecimento de Utilidades

Este módulo filtra as séries pertinentes ao disparo de um alarme tendo como base a análise dos subconjuntos obtidos no módulo “Extração de Médias” e permite, assim, a redução da dimensionalidade dos dados a serem inseridos na etapa de classificação presente no módulo de previsão.

Neste trabalho propõe-se de forma original um método para a escolha do subconjunto de sensores associados a um determinado alarme a partir da definição do fator de utilidade. Tal fator tem como finalidade identificar o grau de importância da variável para o disparo de um alarme, sendo dado por:

$$\varphi = \frac{\mu_{Amax}}{\sigma_t} \quad (4)$$

onde  $Amax$  é a amplitude do maior gradiente entre as médias de regimes contíguos de cada disparo  $d$ , a média destas amplitudes máximas é definida por

$$\mu_{A_{\max}} = \frac{\sum_{d=1}^d A_{\max_d}}{d} \quad (5)$$

e  $\sigma_t$  é o desvio-padrão do instante de tempo em que ocorreram as amplitudes de maior gradiente de cada disparo  $d$ .

O estabelecimento de um limiar explicitará quais variáveis apresentam fortes possibilidades de serem as responsáveis pelo disparo de um determinado alarme, uma vez que um fator de utilidade maior que o limiar implicará a utilização da variável para a predição daquele alarme.

Desta forma, torna-se necessária a normalização das médias das amplitudes máximas de cada variável, a qual é dada por

$$\mu_{A_{\max \text{ norm}}} = \frac{\mu_{A_{\max}} - \min}{\max - \min} \quad (6)$$

onde  $\max$  e  $\min$  são, respectivamente, os valores máximos e mínimos da variável para o período analisado.

Esta metodologia será aplicada a todas as variáveis monitoradas pelo sistema supervisor.

### 3.3 – Previsão de Alarmes

Este módulo tem por objetivo levantar as assinaturas das condições normais e anormais de operação da planta e apresentar diagnóstico e prognóstico a partir da condição vigente de operação.

Classificadores estatísticos consideram que os padrões das classes formam uma população estatística caracterizada por uma determinada função de distribuição ou função de densidade. Estas funções têm, em geral, uma forma desconhecida. Assim, os classificadores tentam ajustá-los aos diferentes modelos teóricos conhecidos, seguindo as abordagens paramétricas ou não paramétricas [8].

O classificador não-supervisionado k-means será utilizado neste trabalho para o reconhecimento de padrões relativos às condições de operação normal e anormal do equipamento com o objetivo de permitir a avaliação da saúde deste. Nesta

técnica de agrupamento os dados são particionados em grupos, cujos integrantes devem ser o mais semelhantes entre si e os dados que pertencem a grupos diferentes devem apresentar a menor semelhança possível [1]. O conjunto de variáveis (sensores) selecionado pelo módulo “Fornecimento de Utilidades” será a entrada para o classificador.

#### 4 – RESULTADOS E DISCUSSÃO

O sistema supervisorio da usina hidrelétrica monitora 166 variáveis e aciona um conjunto de 117 alarmes.

Para fins de validação foi escolhido um dos alarmes para ter seu diagnóstico realizado, cuja escolha se deu em função de sua relevância para o processo de geração de energia elétrica.

Na execução do método ART foi considerada uma janela temporal constante para todos os disparos do alarme e correspondente a um intervalo de tempo de 9 horas. Esta janela foi escolhida em função da dinâmica da planta. O comportamento de uma das variáveis antes da ocorrência de um dos disparos do alarme está apresentado na Figura 2. O momento do alarme encontra-se próximo ao limite da direita do gráfico.

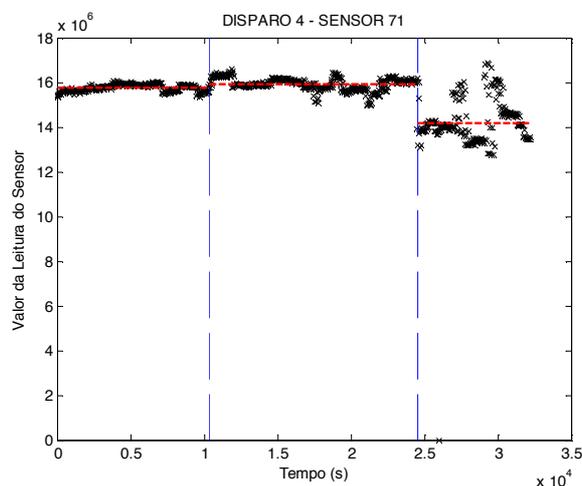


FIGURA 2 – Aplicação do Método ART para uma das Variáveis Monitoradas (sensor 71). As quebras estruturais da série temporal estão representadas pela linha tracejada azul, ao passo que a linha tracejada vermelha representa os valores médios dos regimes de operação (subconjuntos formados) da variável.

Esta metodologia foi aplicada para todas as setenta e três variáveis e envolveu todos os disparos deste alarme. Com os resultados do método ART foram calculados os setenta e três fatores de utilidade da amostra, conforme abordado em “Fornecimento de Utilidades”, item 3.2. A normalização destes fatores levou ao gráfico apresentado a seguir na Figura 3. O gráfico permite observar que apenas dezenove variáveis apresentam fator de utilidade acima do limiar, fato que reduziu drasticamente (redução na ordem de 74%) o número de variáveis a serem consideradas para a classificação.

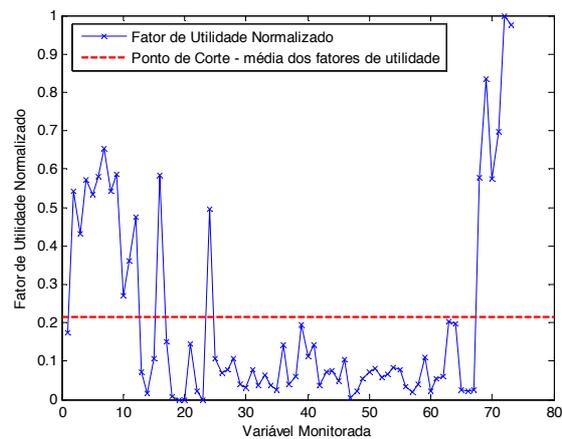


FIGURA 3 – Fornecimento de Utilidades Normalizadas para o Alarme

Uma fusão no nível de dados foi aplicada de forma que a matriz de entrada ( $m \times n$ ) para o classificador apresentou dimensão (1075x19), onde  $m$  corresponde ao número de leituras realizadas na janela temporal anterior ao disparo do alarme e  $n$  a quantidade de sensores selecionados para o referido alarme. O classificador k-means foi realizado com número de clusters pré-definidos e igual a dois. Observa-se pela classificação na Figura 4 que o Cluster 2 representa a condição normal de operação da planta, enquanto que o outro a condição anormal.

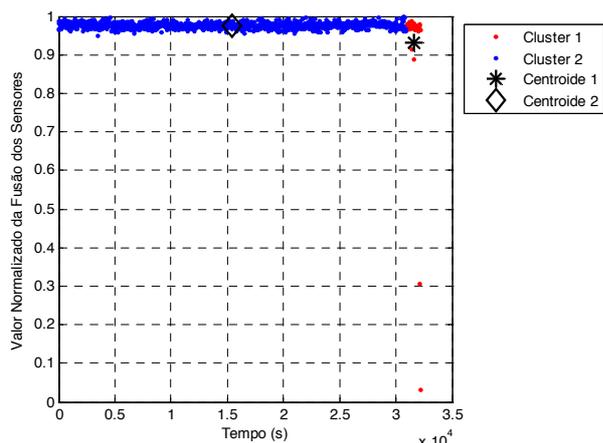


FIGURA 4 – Classificação k-means para dois clusters. Cada ponto possui 19 dimensões, uma vez que houve a fusão das variáveis selecionadas antes da classificação. Nesta figura está representada, ao longo do tempo, apenas uma das 19 dimensões. O momento do alarme encontra-se próximo ao limite da direita do gráfico.

A avaliação da saúde do equipamento e predição de alarmes está sendo desenvolvida a partir do estudo de técnicas de sobreposição estatística entre as assinaturas de comportamento normal e anormal relatadas.

## 5 – CONCLUSÕES

A metodologia empregada na amostra dos dados foi a utilização da Atheoretical Regression Trees (ART), a qual faz uso de uma covariável artificial que é uma seqüência arbitrária de números em ordem crescente ou decrescente, onde esta covariável não é um preditor variável, mas sim um contador. O uso de tal covariável na árvore de regressão de mínimo quadrado recorre a um uso seqüencial do método de Fisher de exata otimização (1958) proposto para o agrupamento de  $n$  elementos em  $g$  subconjuntos com máxima homogeneidade, ou seja, minimiza grupos com a soma dos quadrados.

A aplicação do método ART neste problema foi motivada pela necessidade da busca, para os  $n$  disparos do alarme, de um comportamento padrão das variáveis antes da ocorrência do mesmo.

Fatores de utilidade foram calculados para todas as variáveis presentes na amostra e o estabelecimento de um limiar em função da média deste fator permitiu a

seleção daquelas que apresentam maior probabilidade de uma relação direta com o disparo de um determinado alarme.

Reconhecimento de padrões mediante utilização do método k-means foi aplicado para a classificação de grupos referentes à assinatura de comportamento normal e anormal de operação de um processo industrial.

Os resultados até o momento obtidos pela metodologia apresentada serão validados quando concluído o módulo “Previsão de Alarmes”, item 3.3, que está em fase de desenvolvimento.

## REFERÊNCIAS BIBLIOGRÁFICAS

- [1] BONVENTI JR., W.; COSTA, A. H. R. Sistema semi-automático de detecção de pele por agrupamentos nebulosos. In: VI SIMPÓSIO BRASILEIRO DE AUTOMAÇÃO INTELIGENTE. *Anais...*Bauru, 2003.
- [2] BRAUNSTEIN, S. H.; WERHLI, A. V.; BOTELHO, S. S. C. Análise exploratória de dados para o desenvolvimento de sistema de previsão de alarmes em plantas industriais. In: IX SIMPÓSIO BRASILEIRO DE AUTOMAÇÃO INTELIGENTE. *Anais...* Brasília, set. 2009.
- [3] CAPPELLI, C.; PENNY, R. N.; REALE, M. Detecting multiple mean breaks at unknown points with atheoretical regression trees. em *International Congress on Modelling and Simulation - Advances and Applications for Management and Decision Making*, 2005, pp. 974-978.
- [4] CASOETTO, N.; DJURDJANOVIC, D.; MAYOR, R.; NI, J.; LEE, J. Multisensor process performance assessment through use of autoregressive modeling and feature maps. *Journal of Manufacturing Systems*, vol. 22, pp. 64-72, Dec. 2003.
- [5] DJURDJANOVIC, D.; LEE, J.; NI, J. Watchdog Agent — an infotronics-based prognostics approach for product performance degradation assessment and prediction. *Advanced Engineering Informatics*, vol. 17, pp. 109–125, Jul.- Oct. 2003.
- [6] FACELI, K. *Combinação de métodos de inteligência artificial para fusão de sensores*. Dissertação [Mestrado] - Universidade de São Paulo, São Carlos, Brasil, Fev. 2001.
- [7] FU, C.; YE, L.; LIU, Y.; YU, R.; IUNG, B.; CHENG, Y.; ZENG, Y. Predictive maintenance in intelligent-control-maintenance-management system for

- hydroelectric generating unit. *IEEE Transactions on Energy Conversion*, vol. 19, pp. 179-186, Mar. 2004.
- [8] GIL, J. L.; GARCIA, E. B.; PONVERT-DELISLES, D. R.; SÁNCHEZ, R.; VEJA, M. B. Enfoques para la clasificación digital de imágenes mono y multiespectrales y su implementación en el software cubano TN Estudio V2.0. *Revista de Teledetección*, vol. 20, pp. 35-52, Dic. 2003.
- [9] JARDINE, A. K. S.; LIN, D.; BANJEVIC, D. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, vol. 20, pp. 1483–1510, Oct. 2006.
- [10] LEE, J.; NI, J.; DJURDJANOVIC, D.; QIU, H.; LIAO, H. Intelligent prognostics tools and e-maintenance. *Computers in Industry*, vol. 57, pp. 476–489, Aug. 2006.
- [11] MORA, J.; CORMANE, J.; ORDÓNEZ, G. K-means algorithm and mixture distributions for locating faults in power systems. *Electric Power Systems Research*, vol. 79, pp.714-721, May. 2009.
- [12] *Open Systems Architecture for Condition-based Maintenance (OSA-CBM)*, Machinery Information Management Open Standards Alliance (MIMOSA), 2006.
- [13] TONON JR., W. *Reconhecimento de Símbolos Elétricos*. Dissertação [Mestrado] - Pontifícia Universidade Católica do Paraná, Curitiba, Brasil, 2004.
- [14] WANG, H.F.; WANG, J. P. Fault diagnosis theory: method and application based on multisensor data fusion. *Journal of Testing and Evaluation*, vol. 28, pp. 513–518, Nov. 2000.
- [15] (2009) Electronic Statistics Textbook website. [Online]. Disponível em: <http://statsoft.com/textbook/stathome.html>
- [16] WU, S.; GEBRAEEL, N.; LAWLEY, M. A.; YIH, Y. A neural network integrated decision support system for condition-based optimal predictive maintenance policy. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*, vol. 37, pp. 226-236, Mar. 2007.

## **AGRADECIMENTOS**

Os autores gostariam de agradecer a AES Tietê financiadora do projeto pela base de dados utilizada neste trabalho.