Um modelo neural para a verificação da ocorrência de defeitos em equipamentos eletrônicos com o uso de *Data Mining*, aplicado à área de automação comercial.

Arnaldo Brás de Araujo[1]

[1]Unipac – Universidade Presidente Antônio Carlos
Cep – 36.204-183 – Barbacena – MG – Brazil

arnaldobras@gmail.com

Abstract. *The amount of data available has increased alarmingly in recent years, several factors contributed to this incredible surge. The low cost of storage can be seen as the main cause of the emergence of these huge databases. Another factor is the availability of high-performance computers at a reasonable cost. As a result, these databases now contain treasure troves of information and, due to their size, beyond the technical skills and human capacity in its interpretation there are several alternatives proposed in the literature on how to treat these databases, including KDD and Data Mining. This article proposes and describes a tool to assist in various stages of KDD.*

Resumo. *A quantidade de dados disponíveis vem crescendo assustadoramente nos últimos anos e vários fatores contribuíram para este incrível aumento. O baixo custo na armazenagem pode ser vista como a principal causa do surgimento destas enormes bases de dados. Um outro fator é a disponibilidade de computadores de alto desempenho a um custo razoável. Como consequência, estes bancos de dados passam a conter verdadeiros tesouros de informação e, devido ao seu volume, ultrapassam a habilidade técnica e a capacidade humana na sua interpretação Existem várias alternativas propostas na literatura de como tratar estas bases de dados, entre elas KDD e Data Mining. Este artigo propõe e descreve uma ferramenta para auxiliar em várias etapas do KDD.*

1. Introdução

A constante evolução na iteração homem/maquina fez crescer, de forma inimaginável até pouco tempo, a base de dados armazenada pelos sistemas computacionais. Estes dados por si só não são capazes de agregar valores às organizações a que pertencem. Um dos grandes desafios dos profissionais de TI consiste na melhoria do tratamento destas informações, quer seja no armazenamento, na seleção ou na distribuição. A utilização de sistemas de Business Intelligence permite combinar a recolha e armazenamento de dados com ferramentas de análise, com o objetivo

principal de disponibilizar informação relevante para a tomada de decisão [1]. Estes sistemas são formados por *Data Warehouses* (armazém de dados). Neste contexto surge a Descoberta de Conhecimento em Bases de Dados com o objetivo de desenvolver métodos e técnicas que, a partir de informação guardada em bases de dados, consigam descobrir novo conhecimento, novos padrões e novas tendências, tendo a utilidade e a novidade como critérios de validação [3] [2].

O processo de Descoberta de Conhecimento em Bases de Dados inclui diversas etapas, tais como o pré-processamento e o *Data Mining*, sendo esta última etapa deveras importante, pois é onde efetivamente se procuram as relações entre os dados [4]. Seguindo este principio o uso de *Data Mining* na área de automação comercial pode ser aplicado na predição de fim de uso de equipamentos ou mesmo na descoberta de problemas futuros, que um determinado equipamento pode vir a apresentar.

As mais diversas técnicas empregadas para realizar descobertas de conhecimento, que foram desenvolvidas, sobretudo, pela comunidade de inteligência artificial, tentam encontrar, de maneira automática, regras e modelos estatísticos, que permitem, entre outras coisas, avaliar o comportamento dos dados. De maneira geral, o campo de investigação associado com *Data Mining* combina idéias de descoberta de conhecimento com a implementação eficiente de técnicas, que possibilitam usá-las em banco de dados muito grandes (Silberschatz, 1999) [15].

Em termos concretos, as técnicas de mineração de dados estão relacionadas com o uso de algoritmos, que modelam relações ou padrões não-aleatórios (estatisticamente significativos) em grandes bases de dados (Berry e Linoff, citados por Passari, 2003). Nessa mesma linha de pensamento, Romão et al. (2005) ressaltam que as técnicas de *Data Mining* utilizam dados históricos para aprendizagem, cujo objetivo é realizar uma determinada tarefa particular. Como essa tarefa tem como meta responder alguma pergunta específica de interesse do usuário, é necessário informar qual problema se deseja resolver [15].

É importante ressaltar que não existe um método de mineração de dados universal, portanto a escolha de um algoritmo particular para uma aplicação é de certa forma, uma arte (Fayyad et al., citados por Romão et al., 2005). Nessa mesma perspectiva, a partir da leitura dos trabalhos de Gargano e Raggad (1999) e Berry e Linoff (citados por Passari, 2003), podem ser destacados alguns critérios utilizados para a avaliação e a escolha da técnica mais adequada para atingir um determinado objetivo: robustez, grau de automação, velocidade, poder explanatório, acurácia, quantidade de pré-processamento necessário, escalabilidade, facilidade de integração, habilidade para lidar com muitos atributos, facilidade de compreensão do modelo, facilidade de treinamento, facilidade de aplicação, capacidade de generalização, utilidade e disponibilidade. Ainda de acordo com os referidos autores, os principais fatores que determinam a escolha da técnica a ser utilizada estão relacionados com preponderância de variáveis categóricas ou numéricas, números de campos, número de variáveis dependentes, orientação no tempo e presença de dados textuais [15].

Conclui-se, portanto, que não há critérios universais aplicáveis para a escolha e utilização de técnicas de mineração de dados. Isso porque cada técnica possui critérios específicos que devem ser levados em consideração. assim, de acordo com Passari

(2003), é também extremamente difícil comparar as técnicas entre si, já que operam de maneira distinta. O autor conclui que a única forma de avaliá-las é por meio da medição de sua habilidade em desempenhar as tarefas para a quais foram construídas [15].

Apesar de cada técnica de mineração de dados ter sua própria abordagem, elas compartilham algumas características em comum: conforme "aprendem" a partir dos dados de treinamento coletados, ela melhora, gradativamente, a sua performance; e existe sempre uma fase de treinamento, onde o modelo "aprende" os padrões e os relacionamentos (essa fase de treinamento é seguida pela fase implementação, quando o modelo é posto à prova) (Passari, 2003) [15].

Para encontrar respostas, ou extrair conhecimentos interessantes, existem diversas técnicas de mineração de dados. A partir da leitura de alguns trabalhos (Bispo, 1998; Elmasri e Navathe, 2002; Passari, 2003; Wong e Leung, 2002; Romão et al., 2002) que enfocam esse tema, pode-se citar seis técnicas principais relacionadas com *Data Mining*: indução de regras, redes neurais, algoritmos genéticos, árvores de regressão, lógica nebulosa e clustering [15].

Técnicas de **indução de regras** consistem no uso de ferramental matemático e estatístico, que visam o desenvolvimento de relacionamentos a partir dos dados apresentados. Tipicamente, são criadas correspondências do tipo "se-então", baseadas em relações causais detectadas nas variáveis. Cada relacionamento "se-então" extraído é chamado de "regra" [15].

As **redes neurais** são técnicas derivadas de pesquisas, na área da inteligência artificial, que utilizam à regressão generalizada. Essas técnicas fornecem "métodos de aprendizagem", pois são conduzidas a partir de amostragens de testes, utilizadas para inferências e aprendizagem iniciais. Com esses métodos de aprendizagem, respostas a novas entradas podem ser passíveis de serem interpoladas a partir das amostras conhecidas. Essa interpolação, no entanto, depende do modelo mundial desenvolvido através do método de aprendizagem [15].

Os **algoritmos genéticos** estão relacionados com técnicas de otimização, onde se utilizam combinações de processos (exemplo: combinação genética, mutação e seleção natural). Essas técnicas estão associadas, sobretudo, com o conceito de seleção natural [15].

As **árvores de regressão** são técnicas simples, baseadas na autonomia de uma árvore. Nesse sentido, cada galho particiona, de forma estratégica e sucessiva, os dados em classes e subclasses. A cada divisão, é escolhida a melhor forma de separar e classificar os dados, de acordo com a característica que mais os distingue. Para isso, são utilizadas, também, medidas estatísticas. As árvores de regressão possuem algoritmos não-supervisionados, ou seja, são capazes de processar automaticamente os dados [15].

Técnicas de **lógica nebulosa** são utilizadas para capturar informações vagas, que, em geral, são descritas na sua forma natural, e convertê-las em um formato numérico, para facilitar as suas análises. Em termos operacionais, essas técnicas utilizam a teoria dos conjuntos nebulosos, que tem mostrado ser muito apropriada para se trabalhar com vários tipos de dados e informações, superando, muitas vezes, os

resultados obtidos com o emprego das tradicionais técnicas estatísticas e probabilísticas [15].

O **clustering** está relacionado com técnicas de *Data Mining* direcionadas aos objetivos de identificação e classificação. Essa técnica tenta identificar um conjunto finito de categorias, ou clusters, para os quais cada objeto de dado pode ser mapeado. As categorias podem ser disjuntas ou sobrepostas e, às vezes, ser organizadas em árvores [15].

Esta ferramenta possui diversas aplicações em inúmeras áreas tais como: área médica (utilizado na predição de possíveis portadores de câncer, no estudo de históricos cirúrgicos para descobrir fatores que afetam o sucesso ou fracasso de cirurgias na coluna), na área de ciências e tecnologias (pesquisas com estruturas moleculares, dados genéticos, mudança no clima, etc), na área de marketing (descoberta de clientes de risco, montar relação de produtos mais vendidos, etc), na área financeira (descoberta de fatores de risco para aplicações e investimentos, etc) [13] [10].

Por este motivo os profissionais da área tem empenhado grandes esforços para se conseguir alcançar bons resultados, tendo em vista o enorme beneficio que isto pode trazer para as organizações. Porém as dificuldades encontradas também são enormes devido à complexidade dos dados e ao enorme volume de dados armazenados. Outro fator importante é o desafio da preparação dos dados para a mineração pois envolvem tarefas trabalhosas e inúmeras, sendo consideras como 80% do trabalho da descoberta de conhecimento em uma base de dados [10].

Diante de tais problemas, e tendo em vista a grande quantidade de dados armazenados pelas empresas e o quanto que um conhecimento especifico desta base de dados pode ser importante para as organizações, trabalhos realizados nesta área podem vir a contribuir para uma evolução e aprimoramento das técnicas existentes, ou mesmo o surgimento de novas técnicas [10].

Com isto, o objetivo deste trabalho é propor um modelo que a partir da análise de uma base de dados seja capaz de prever os possíveis problemas que um determinado equipamento eletrônico pode vir a apresentar e quando estes problemas poderão ocorrer. A partir do estudo e avaliação da literatura, será utilizada a técnica de redes neurais artificiais (RNAs) do tipo multy-layer perceptrons (MLPs) supervisionada e treinada pelo algoritmo backpropagation.

2. Metodologia

A metodologia utilizada no processo de *Data Mining* será a CRISP-DM (Cross Industry Standard Process for *Data Mining*). A metodologia CRISP-DM foi desenvolvida por um consórcio composto por NCR Systems Engineering Copenhagen, DaimlerChrysler AG, SPSS Inc. e OHRA Verzekeringen en Bank Groep B.V [5].

No guia publicado pelo consórcio [5] o processo de *Data Mining* é encarado como um projeto com um ciclo de vida iterativo, tal como se pode observar na Figura 1. Esse ciclo de vida consiste em seis fases cuja sequência não é rígida, mas sim dependente do resultado de cada fase. Na Figura 1 representam-se com setas as relações mais frequentes entre as várias fases.

A primeira parte consiste na **compreensão do negócio**, definição do problema de *Data Mining* e criação de um plano preliminar. Em seguida vem a parte de **compreensão dos dados**, onde se identificam problemas nos dados, subconjuntos de dados e formula-se hipótese. A terceira fase consiste na **preparação dos dados** onde serão produzidos os conjuntos de dados (*dataset*). A fase seguinte é a fase de **modelação** onde será aplicada a técnica de *Data Mining*, calibrada de forma a obter os melhores resultados. Neste trabalho será usada a técnica de redes neurais artificiais treinada pelo algoritmo de backpropagation. A quinta fase consiste na **avaliação dos resultados** da fase anterior. Nesta fase também é feita uma revisão de todos os passos anteriores para garantir que os resultados obtidos cumprem o objetivo do negócio. A ultima fase, é a fase do **desenvolvimento** que consiste na produção de relatórios e na repetição de todo o processo de *Data Mining*. Esta parte é realizada pelo cliente, no entanto o analista deve certificar-se que o cliente compreendeu todas as ações para utilização dos modelos criados [5].

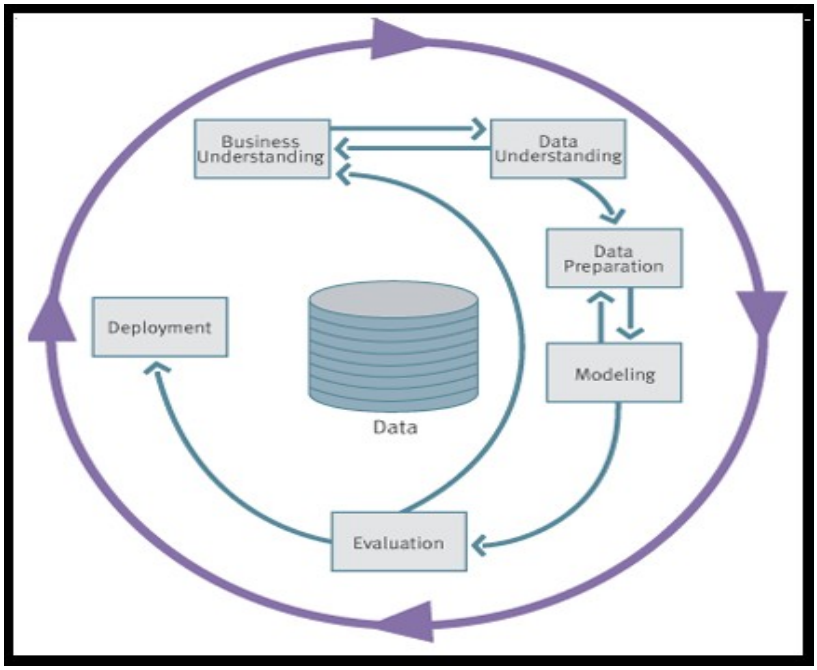


Figura 1 - Metodologia CRISP-DM

Fonte: **[5]**

2.1 Compreensão do négocio

A fase de compreensão do negocio foi realizada a partir do conhecimento especialista adquirido ao longo de um período de três anos de trabalho de manutenção em equipamentos eletrônicos.

2.2 Compreensão dos dados

A fase de compreensão dos dados, assim como do negocio, foi realizada a partir do conhecimento especialista adquirido ao longo de um período de três anos de trabalho com equipamentos eletrônicos.

2.3 Preparação dos dados

A partir de uma base de dados coletada manualmente ao longo do trabalho, foram separados dados relativos ao histórico de manutenção corretiva em balanças platina 15 kg. A partir daí estes dados foram transferidos para uma planilha eletrônica, que em seguida foram salvos no formato CSV. A partir do arquivo CSV, foi gerado um arquivo. ALL, o qual foi importado para o KNIME.

2.4 Modelação

2.4.1. Redes Neurais Artficiais

A disciplina de Inteligência Artificial almeja o desenvolvimento de paradigmas computáveis capazes de realizar tarefas cognitivas que só são realizadas pelo cérebro humano [6]. A partir daí começou a surgir as redes neurais artificiais (RNAs). O termo redes neurais artificiais tende a criar grandes expectativas. Entretanto, vale salientar que o modelo de rede neural é um modelo simplificado do sistema nervoso central dos seres humanos, passivo de implementação por software ou hardware o qual é capaz de realizar tarefas (como classificação, regressão e predição) se submetida a um período de treino [8].

As redes RNAs evoluíram a partir do perceptron, proposto por Rosenblatt 1962[7], o qual é formado por um conjunto de neurônios artificiais e sinapses que os interligam. Cada neurônio calcula a média dos sinais que recebe e a armazena nas respectivas sinapses. Este valor é usado como entrada de uma função de ativação e o resultado desta função é o valor de saída do neurônio, o qual pode ser ou não usado por outro neurônio [8].

A topologia das redes RNAs se caracterizam no modo como os neurônios se interligam. É condicionada ao fato do aprendizado da rede poder ser supervisionado ou não. Uma rede supervisionada é constituída basicamente por três camadas de neurônios. Uma camada de entrada, uma camada intermediaria e uma camada de saída. Na aprendizagem supervisionada o processo de aprendizagem ou treino da rede consiste em ajustar os valores dos pesos das sinapses de modo que as entradas na rede produzam as saídas corretas [8] [11].

As redes do tipo MLP são constituídas da mesma forma das outras RNAs, porem a topologia da rede é outra. Redes do tipo MLP possuem topologia em camadas e as conexões são unidirecionais [8].

2.4.2. Aprendizagem

Aprendizagem é o processo pelo qual os parâmetros de uma rede neural são adaptados por meio de um estímulo contínuo do ambiente, no qual a rede está operando, sendo o tipo específico de aprendizagem realizada definido pela maneira particular como ocorrem os ajustes realizados nos parâmetros (Haykin, 1994) [14] [12].

Sem dúvida, o grande atrativo das redes neurais está na capacidade de aprendizagem, onde a rede extrai informações relevantes de padrões apresentados gerando uma representação própria do problema [14].

O processo de aprendizagem pode ser simplificado da seguinte forma (Trippi; Turban, 1993): “As Redes Neurais Artificiais aprendem com seus erros”. Normalmente o processo de aprendizagem (ou treinamento) envolve três fases [14]:

1- Compute a saída;

2- Compare a saída com as respostas desejadas;

3- Ajuste os pesos e repita o processo;

Usualmente o processo de aprendizado começa pelos pesos randomizados. A diferença entre a saída atual (Y ou Yt) e a saída desejada (Z) é chamada de Δ. O objetivo é minimizar o Δ (ou no melhor reduzi-lo a zero). A redução do Δ é feita pela mudança incremental dos pesos “[14]”.

2.4.3. Algoritmo Backpropagation

No algoritmo backprogation o processo de aprendizagem é dividido em ciclos. O objetivo é obter o mínimo de erro nas saídas geradas. Para isto, os valores obtidos pela diferença entre os valores de saída da rede e os valores de treino são jogados para trás desde os neurônios de saída até os neurônios de entrada, e em seguida o peso das conexões é reajustado. O algoritmo termina quando o mínimo de erro à saída for obtido. Este algoritmo é computacionalmente pesado e de convergência lenta, e algumas melhorias já foram propostas tais como o QuickPropagation desenvolvido por Fahlman em 1998, ou o RPPROP (Resilient Backpropagation) proposto por Riedmiller. Entretanto, o algoritmo backprogation, permite uma forma automática de ajuste do peso das sinapses. Os outros algoritmos mencionados são uma evolução do Backprogation. Desta forma o estudo deste algoritmo possui uma relevância maior no estudo do processo de aprendizagem [14].

O algoritmo “Backpropagation” (BP) refere-se a uma regra de aprendizagem que consiste no ajuste dos pesos e polarizações da rede através da retro propagação do erro encontrado na saída. A minimização é conseguida realizando-se continuamente – a cada interação – a atualização dos pesos e das polarizações da rede, no sentido oposto ao do gradiente da função no ponto corrente, ou seja, proporcionalmente ao negativo da derivada do erro quadrático em relação aos pesos correntes. Trata-se, portanto, de um algoritmo de treinamento supervisionado, determinístico, de computação local, e que implementa o método do gradiente decrescente nas somas dos quadrados dos erros [14].

A topologia da arquitetura da rede que utiliza esta regra de aprendizagem é formada, geralmente, por uma ou mais camadas escondidas (intermediárias) de neurônios não-lineares (com função de propagação sigmoidal) e uma camada de saída de neurônios lineares. Devido a grande difusão da arquitetura da rede a que esta regra de aprendizagem se aplica, é comum referir-se a ela com o nome da própria regra de aprendizagem, ou seja, rede BP. Redes BP com polarizações com, no mínimo, uma camada intermediária de neurônios lineares na saída, são teoricamente capazes de realizar a aproximação de qualquer função matemática, sendo ainda bastante utilizadas na associação e classificação de padrões [14].

3. Resultados

3.1 Pré-seleção das variáveis

As variáveis foram pré-selecionadas através de conhecimento especialista, e estão descritas conforme as tabelas 1.0 e 1.1 a seguir. Esta variável compõe a base de dados com a lista dos problemas encontrados, juntamente com as soluções encontradas. São elas:

TABELA PROBLEMA		
Código	***Descrição***	***Cod.***
1	Impressão ruim	0.1
2	Não comunica	0.2
3	Desligando	0.3
4	Display quebrado	0.4
5	Peso errado	0.5
6	Teclado ruim	0.6
7	Desconfigurada	0.7
8	Não liga	0.8
9	Não imprime	0.9

Tabela 1.0 – Tabela com as variáveis problemas

TABELA SOLUÇÃO		
Código	***Descrição***	***Cod.***
1	Regulagem impressor	0.1
2	Reparo placa principal	0.2
3	Troca do display	0.3
4	Calibração	0.4
5	Troca do teclado	0.5
6	Reparo ponto de rede	0.6
7	Troca cabeça de impressão	0.7
8	Reparo placa fonte	0.8
9	Reparo cabo de energia	0.9

Tabela 1.1 – Tabela com as variáveis solução

Estas foram variáveis utilizadas na primeira parte do problema que foi a implantação do modelo de predição de problemas e prováveis soluções.

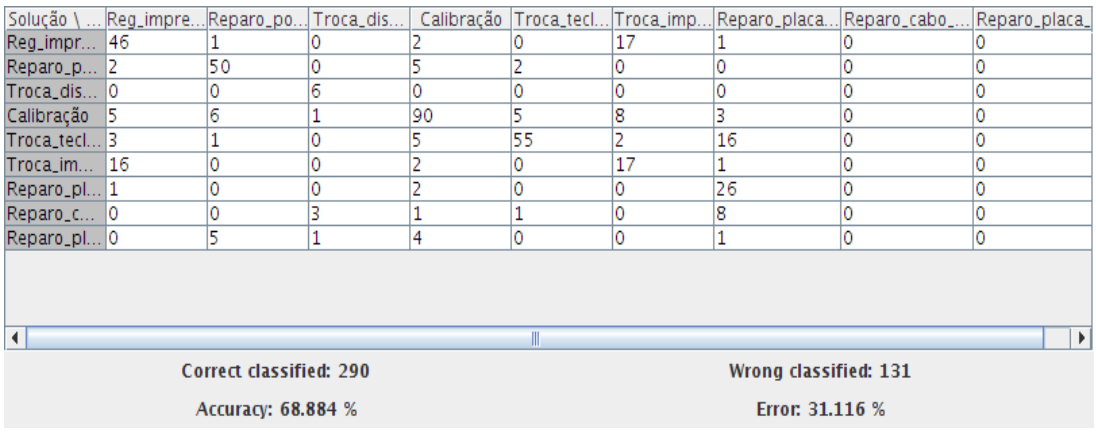
3.2 Coleta de dados

Foram coletados 1.400 pontos de operação para o processo de treinamento da rede neural. O tratamento dos dados foi baseado em: (i) Análise dos dados a fim de eliminar redundância para evitar que a rede se torne tendenciosa; (ii) Eliminação de valores impróprios; (iii) Classificação e separação dos defeitos apresentados ao longo do período, bem como as devidas soluções encontradas para tais problemas.

3.3 Treinamento da rede neural

A rede de melhor desempenho foi a MLP com cinco camadas intermediárias, ou camadas ocultas, e com 10 neurônios por camada. O treinamento foi realizado utilizando o algoritmo backpropagation. É um algoritmo supervisionado que utiliza pares (entrada, saída desejada) para, por meio de um mecanismo de correção de erros, ajustar os pesos da rede. O treinamento ocorre em duas fases, em que cada fase percorre a rede em um sentido. Estas duas fases são chamadas de fase forward e fase backward. A fase forward é utilizada para definir a saída da rede para um dado padrão de entrada. A fase backward utiliza a saída desejada e a saída fornecida pela rede para atualizar os pesos de suas conexões.

Nos testes realizados com a primeira base, para descobrir os problemas que mais ocorrem e predizer problemas futuros, os valores são mostrados na figura 2.0.



Solução \ ...	Reg_impre...	Reparo_po...	Troca_dis...	Calibração	Troca_tecl...	Troca_imp...	Reparo_placa...	Reparo_cabo_...	Reparo_placa_
Reg_impr...	46	1	0	2	0	17	1	0	0
Reparo_p...	2	50	0	5	2	0	0	0	0
Troca_dis...	0	0	6	0	0	0	0	0	0
Calibração	5	6	1	90	5	8	3	0	0
Troca_tecl...	3	1	0	5	55	2	16	0	0
Troca_im...	16	0	0	2	0	17	1	0	0
Reparo_pl...	1	0	0	2	0	0	26	0	0
Reparo_c...	0	0	3	1	1	0	8	0	0
Reparo_pl...	0	5	1	4	0	0	1	0	0

Correct classified: 290 Wrong classified: 131

Accuracy: 68.884 % Error: 31.116 %

Figura 2.0 – Matriz confusão primeira base.

Com os valores obtidos, é possível avaliar o desempenho do modelo. Para isto foram realizados diversos testes, e calculada a média dos valores obtidos, conforme pode ser visto na tabela 1.2. Os valores, basicamente, nos mostram a imprecisão do modelo para predizer o comportamento durante o processo.

	TESTE I	TESTE II	TESTE III	TESTE IV	TESTE V	TESTE VI	Média	Desvio padrão
Acurácia	67,3	67,39	69,21	68,11	69,74	69	68,46	1,01
Precisão	56,04	54,58	56,06	58,74	58,66	58,19	57,04	1,72
Sensibilidade	28,69	30,55	30,49	32,5	30,38	33,56	31,03	1,73
Especialidade	65,35	69,41	63,29	71,23	69,77	70,35	68,23	3,16
Acurácia ajustada	47,02	49,98	46,89	51,87	50,08	51,95	49,63	2,24

Tabela 1.2 – Tabela com a média dos valores obtidos.

4. Trabalhos futuros

O processo de *Data Mining* é um processo extenso. A partir dos resultados obtidos é necessário extrai da base de dados obtida o conhecimento por ela gerado. As redes neurais artificiais, apesar de eficientes, deixam a desejar neste ponto, pois não mostram o caminho percorrido para se chagar até o conhecimento obtido. Para isto é necessário o auxilio de técnicas que possam extrair este conhecimento. Dentre as técnicas mais utilizadas, destaca-se o algoritmo Trepan.

O algoritmo Trepan foi desenvolvido por Craven e Shavlik [CS96]. Este algoritmo faz perguntas a uma RNA treinada utilizando os padrões de entrada do seu conjunto de treinamento. As respostas a estas perguntas são utilizadas para a construção de uma árvore de decisão que aproxima o conhecimento representado pela rede. Trepan aborda a extração de um conjunto compreensível de regras de uma rede neural treinada como sendo um problema de aprendizado indutivo. Neste aprendizado, o objetivo é a função representada pela rede, e a descrição do conhecimento adquirido pela rede é uma árvore de decisão que a aproxima. Para evitar que as árvores extraídas fiquem muito grandes, Trepan utiliza um parâmetro para limitar o número de seus nós internos.

Para guiar a construção da árvore, Trepan utiliza um oráculo, capaz de responder a perguntas durante o processo de construção. O oráculo, por sua vez, faz perguntas à rede, uma vez que é na rede que está presente o conhecimento a ser extraído. A vantagem de aprender por perguntas, em vez de por exemplos, é que as perguntas podem conseguir, de uma forma mais precisa e específica, as informações que se mostrarem necessárias.

5. Conclusão

Este trabalho apresentou a formação de um modelo completo de mineração de dados (ou de Descoberta de Conhecimento em Bases de Dados) para a predição de problemas em balanças eletrônicas. Utilizou-se como fonte de estudos, base de dados própria, adquirida ao longo do tempo de trabalho com estes equipamentos, visando à predição da ocorrência dos problemas mais comuns que ocorreram ao longo do tempo.

O processo completo de mineração de dados foi descrito, citando as principais metodologias e ferramentas. Tendo em vista o problema proposto, foi utilizada a técnica de Redes Neurais, treinada e supervisionada pelo algoritmo backpropagation na fase de predição e classificação de problemas.

O modelo desenvolvido utilizou como embasamentos principais o processo de Descoberta de Conhecimentos em Bases de Dados, proposto por Fayyad et al. (1996) e a metodologia CRISP-DM, definida em CRISP (2000).

Tendo em vista as limitações e dificuldades para a formação dos data marts, entende-se que os resultados demonstram viabilidade de aplicação prática do modelo desenvolvido, com potencial de aprimoramento e continuidade. O modelo desenvolvido pode ser estudado e adaptado para outros problemas envolvendo predição e classificação.

- Aperfeiçoamento do conjunto de dados de treinamento, através da inserção de variáveis complementares;
- Formação do conjunto de treinamento variando a técnica de amostragem utilizada, buscando uma diversidade maior para a descoberta de padrões de comportamento;
- Variação do conjunto de algoritmos classificadores, buscando possibilidades de otimização dos resultados;

References

1. KING, D. CS 4803B- "Numerical Machine Learning", 1995.
2. http://www.crisp-dm.org/CRISPWP-0800.pdf (2006)
3. CORTEZ, P., "Modelos inspirados na Natureza Para a Previsão de Séries Temporais", Tese de Doutorado, Universidade do Minho, 2002.
4. SANTOS, M.Y; RAMOS I.; "Business Inteligence", FCA – Editora Informática, 2006.
5. FAYYAD, U.M et al. Advances in Knowledge Discovery and *Data Mining*, M.I.T. Press, 1996.
6. FAYYAD, U.; SHAPIRO G.; SMYTH P. From *Data Mining* to Knowledge Discovery in Databases, AI Magazine, 1996.
7. VAPNIK V. The Nature of Statistical Learning Theory, Springer-Verlag, New York, 1995.
8. BISHOP C. Neural Networks for Pattern Recognition, Oxford Univ. Press, 1995.
9. CRUZ, Armando J.R. *Data Mining* via Redes Neuronais Artificiais e Máquinas Vectores de Suporte, Tese de Mestrado, Universidade do Minho, 2007.
10. SERGION. Princípios Essenciais do *Data Mining*. Disponível em: http://www.intelliwise.com/snavega.
11. BARRETO, Jorge M. Introdução as Redes Neurais Artificiais. Disponível em: http://www.das.ufsc.br/~gb/CIPEEL/Survey.pdf
12. BARRETO, Guilherme A. Redes Neurais Artificiais: Uma Introdução Pratica. Disponível em: http://www.deti.ufc.br/~guilherme/TI016/slides_PS_MLP.pdf

13. http:// www.the-data-mine.com (2006)

14. JUNIOR, Rubião G.T.; MACHADO, MARIA A.S.; SOUZA, Reinaldo C. Previsão de Séries Temporais de Falhas em Manutenção Industrial Usando Redes Neurais. Disponivél em : www.uff.br/engevista/2_7Engevista03.pdf

15. LAZZAROTTO, Lissandra L.; OLIVEIRA, Alcione P.; LAZZAROTTO, Joelsio J. Aspectos teóricos do *Data Mining* e aplicação das redes neurais em previsões de preços agropecuários. UFV.