



PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ
ESCOLA POLITÉCNICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO
E SISTEMAS - PPGEPS

RUY GOMES DA SILVA

REDUÇÃO DOS IMPACTOS AMBIENTAIS OCACIONADOS POR UM
PROCESSO DE PRODUÇÃO DE ALUMINA COM O USO DE
TÉCNICAS DE RECONHECIMENTO DE PADRÕES

CURITIBA/PR
2014

RUY GOMES DA SILVA

**REDUÇÃO DOS IMPACTOS AMBIENTAIS OCACIONADOS POR UM
PROCESSO DE PRODUÇÃO DE ALUMINA COM O USO DE
TÉCNICAS DE RECONHECIMENTO DE PADRÕES**

Projeto de Dissertação apresentado ao Curso do Programa de Pós-Graduação em Engenharia de Produção e Sistemas (PPGEPS) da Pontifícia Universidade Católica do Paraná como parte do requisito parcial para obtenção do título de Mestre em Engenharia de Produção.

Orientadora: Prof^a. Dr^a. Maria
Teresinha Arns Steiner.

**CURITIBA/PR
2014**

RUY GOMES DA SILVA

**REDUÇÃO DOS IMPACTOS AMBIENTAIS OCACIONADOS POR UM
PROCESSO DE PRODUÇÃO DE ALUMINA COM O USO DE
TÉCNICAS DE RECONHECIMENTO DE PADRÕES**

Projeto de Dissertação apresentada ao Curso do Programa de Pós-Graduação em Engenharia de Produção e Sistemas (PPGEPS) da Pontifícia Universidade Católica do Paraná como parte do requisito parcial para obtenção do título de Mestre em Engenharia de Produção.

COMISSÃO EXAMINADORA

Prof^a. Dr^a.Eng. Maria Teresinha Arns Steiner (PUCPR)
Orientadora

Prof. Dr. Sc. Anderson Roges Teixeira Góes (UFPR)
Membro Externo

Prof. Dr.Eng. Rui Francisco Martins Marçal (PUCPR)
Membro Interno

CURITIBA-PR, _____ DE _____ DE 2014

DEDICATÓRIA

As minhas filhas, Ágata e Eva, que me
inspiraram em todos os momentos.

AGRADECIMENTOS

Agradeço, primeiramente, a DEUS, por me fortalecer diariamente durante toda esta caminhada, mostrando-me incansavelmente que valeria a pena não desistir.

Aos meus pais, José Camilo da Silva (em memória) e Ercilha Gomes da Silva, pelos esforços e incentivo para a realização deste sonho e as minhas filhas, Ágata dos Santos Silva e Eva Ryanne dos Santos Silva, por compreenderem minha ausência por longos períodos durante a realização deste trabalho.

A professora Dra. Maria Teresinha Arns Steiner, mais que uma orientadora, uma grande incentivadora e amiga. Pessoa fundamental para realização deste sonho.

Ao Professor Dr. Anderson Roges Teixeira Góes, que me Coorientou sabiamente e me auxiliou em muitos momentos difíceis deste trabalho.

A amiga Tainan Neves, mais que uma simples “colega” de mestrado, uma verdadeira amiga, que desde o início deste projeto fez-se presente nos momentos fundamentais de minha vida.

Ao amigo Lee Vinagre, que também caminhou paralelamente neste projeto, sempre presente nas discussões em nosso laboratório de estudos.

Aos amigos Humberto Rodrigues Lima e Luciano Fernandes de Carvalho, mais que amigos, verdadeiros irmãos, aos quais sempre pude confiar nos momentos difíceis da vida.

Aos amigos de mestrado com quem pude fazer as disciplinas, com os quais sempre estudávamos juntos para superar nossas dificuldades.

Agradeço também a todos os professores com quem tive contato, seja em sala de aula ou numa simples conversa sobre a pesquisa. Tenho certeza que aproveitei cada momento para a consolidação deste trabalho.

RESUMO

O principal processo de extração de Alumina da bauxita foi desenvolvido em 1858 por *Le Chatelier* e aperfeiçoado posteriormente em 1888 por *Karl Bayer*, sendo conhecido atualmente como processo *Bayer*. Cada empresa deste ramo, mesmo que utilize o mesmo processo de extração, tem um diferente valor de concentração de soda cáustica no rejeito, devido melhorias tecnológicas implementadas por cada unidade fabril. Estima-se um valor de 15g/l de soda cáustica para cada tonelada de rejeito gerado sendo que, atualmente, este material é acondicionado em bacias próprias expostas ao meio ambiente. Assim sendo, este trabalho tem por objetivo realizar a previsão deste teor cáustico no rejeito bem como utilizar técnicas de Reconhecimento de Padrões (RP), de forma comparativa, visando reconhecer no rejeito de um processo de produção de alumina, denominado lama vermelha, os diferentes teores de soda cáustica gerada pelas diferentes entradas. Este RP poderá servir de base para estudos futuros, com a possibilidade de construir um modelo matemático capaz de minimizar o teor cáustico contido no rejeito, através da melhor combinação entre variáveis de entrada do processo. Desta forma, o reaproveitamento da lama vermelha poderá ser maior junto às indústrias, como na fabricação de cerâmica e como componente do concreto na indústria civil obtendo-se, como consequência, a minimização dos impactos ambientais gerados pela mesma. As técnicas de RP aqui abordadas (*Redes Neurais Artificiais* e *Support Vector Machine*), abordadas comparativamente visando a classificação e previsão de padrões, consideram os registros históricos referentes às variáveis de entrada e de saída da etapa do processo que realiza a lavagem da lama vermelha. Tais registros são referentes à leitura das medições dos diversos instrumentos envolvidos; a análise dos resultados laboratoriais e do comportamento do filtro rotativo que realiza a lavagem. As *Redes Neurais Artificiais* demonstram-se com resultados muito eficientes neste estudo, tanto na previsão quanto na classificação dos padrões.

Palavras-chave: *Redes Neurais Artificiais. Support Vector Machine. Alumina. Reconhecimento de Padrões.*

ABSTRACT

The main process of extracting alumina from bauxite was developed in 1858 by Le Chatelier in 1888 and later perfected by Karl Bayer, currently known as the Bayer process. Each branch of this company, even when using the same extraction process has a different value of caustic soda concentration in waste, due to technology improvements implemented by each plant. It is estimated a value of 15g / l of caustic soda per ton of reject generated and, currently, this material is packaged in its own basins exposed to the environment. Therefore, this paper aims to carry out the prediction of this caustic content in the waste and use techniques for Pattern Recognition (PR), in a comparative way, in order to recognize reject a process of alumina production, called red mud, the different levels of caustic soda generated by different inputs. This RP can serve as a basis for future studies, with the possibility of constructing a mathematical model capable of minimizing the caustic level contained in the waste, through the best combination of input variables of the process. Thus, the reuse of red mud may be greater through the industry, such as in the manufacture of ceramics and how the concrete component in obtaining civilian industry, as a consequence, the minimization of environmental impacts generated by the same. The PR techniques covered here (Artificial Neural Networks and Support Vector Machine), dealt comparatively aimed at classifying and predicting patterns, consider the historical records related to the input variables and the output stage of the process that performs the washing of red mud. These records are related to meter reading of the various instruments involved and the analysis of laboratory results and behavior of the rotating filter that performs the wash. Artificial Neural Networks show yourself with very effective results in this study, both in the forecast as the classification of patterns.

Keywords: Artificial Neural Networks. Support Vector Machine. Alumina. Pattern Recognition.

LISTA DE TABELAS

TABELA 1 – ANÁLISE DOS COMPONENTES DA BAUXITA.....	25
TABELA 2 – EXEMPLOS DE CLASSIFICAÇÃO	55
TABELA 3 – DADOS EXEMPLO SVM.....	75
TABELA 4 – DADOS DO PROCESSO DE LAVAGEM DE LAMA	80
TABELA 5 – CLASSIFICAÇÃO BINÁRIA DAS AMOSTRAS	81
TABELA 6 – CRIAÇÃO DOS SUB-GRUPOS.....	82
TABELA 7 – CLASSIFICAÇÃO COM SVM.....	85
TABELA 8 – RESUMO DOS RESULTADOS COM SVM.....	86
TABELA 9 – RESUMO DOS RESULTADOS COM RNA	86
TABELA 10 – ESTRUTURAÇÃO DOS RESULTADOS DOS TESTES	87
TABELA 11 – PARÂMETROS DO TREINAMENTO DA RNA COMO RP.....	88
TABELA 12 – MELHORES TAXAS DE ACERTOS DAS RNA'S COMO RP.....	88
TABELA 13 – ESTRUTURAÇÃO DOS RESULTADOS DOS MSE'S DAS RNA'S COMO PREVISÃO.....	89
TABELA 14 – RESULTADOS DOS MSE'S DAS RNA'S DE CLASSIFICAÇÃO	90
TABELA 15 – RESULTADOS DOS MSE'S DAS RNA'S DE PREVISÃO	91
TABELA 16 – PARÂMETROS DE TREINAMENTO DA RNA COMO PREVISÃO....	92
TABELA 17 – MELHORES RNA'S DESENVOLVIDAS.....	93

LISTA DE FIGURAS

FIGURA 1 – MINÉRIO DE BAUXITA EM SUA FASE DE EXTRAÇÃO.....	24
FIGURA 2 – FLUXOGRAMA DO PROCESSO BAYER.....	26
FIGURA 3 – ALUMINA CALCINADA.....	30
FIGURA 4 – PROCESSO DE LAVAGEM DA LAMA VERMELHA.....	32
FIGURA 5 – ÁREA DE REJEITO DE LAMA VERMELHA / EMPRESA NO ESTADO DO PARÁ- BRASIL.....	33
FIGURA 6 - ÁREA DE REJEITO DE LAMA VERMELHA / EMPRESA NO ESTADO DO MARANHÃO BRASIL.....	34
FIGURA 7 – ÁREA DE REJEITO DE LAMA VERMELHA.....	34
FIGURA 8 – SAÍDA DE LAMA VERMELHA APÓS LAVAGEM.....	35
FIGURA 9 – TIJOLOS (BLOCOS) PRODUZIDOS COM A LAMA VERMELHA.....	36
FIGURA 10 – TIJOLOS (BLOCOS) PRODUZIDOS COM A LAMA VERMELHA.....	37
FIGURA 11 – CASA CONSTRUÍDA COM RESÍDUO DO PROCESSO BAYER.....	37
FIGURA 12 – ETAPAS DO PROCESSO KDD POR FAYYAD ET AL., 1996.....	51
FIGURA 13 – FASES DO RECONHECIMENTO DE PADRÕES.....	56
FIGURA 14 – FLUXOGRAMA DE UM PROCESSO DE PREVISÃO.....	59
FIGURA 15 – REPRESENTAÇÃO DA COMUNICAÇÃO ENTRE NEURÔNIOS BIOLÓGICOS.....	62
FIGURA 16 – MODELO DE UM NEURÔNIO ARTIFICIAL.....	63
FIGURA 17 – RNA COM UMA CAMADA OCULTA.....	65
FIGURA 18 – FLUXOGRAMA DO PROCESSO DE APRENDIZAGEM DA RNA.....	66
FIGURA 19 – REPRESENTAÇÃO DA APRENDIZAGEM SUPERVISIONADA.....	67
FIGURA 20 – CONJUNTO DE DADOS LINEARMENTE SEPARÁVEIS.....	69
FIGURA 21 – POSSIBILIDADES DE SEPARAÇÃO POR MEIO DE DIFERENTES HIPERPLANOS.....	70
FIGURA 22 – SEPARAÇÃO ÓTIMA DE DOIS CONJUNTOS LINEARMENTE SEPARÁVEIS.....	70
FIGURA 23 – VETORES SUPORTES E FUNÇÃO DE DECISÃO.....	71
FIGURA 24 – ESPAÇO DE ENTRADA DOS PADRÕES E ESPAÇO DE CARACTERÍSTICAS.....	72
FIGURA 25 – REPRESENTAÇÃO DO SVM NA ESTRUTURA DE RNA.....	72
FIGURA 26 – ESQUEMA DE TREINAMENTO SUPERVISIONADO.....	73
FIGURA 27 – ESQUEMA DE TESTE.....	74
FIGURA 28 – MÉTODO K-FOLDS (K=3).....	84

LISTA DE GRÁFICOS

GRÁFICO 1 – GRÁFICO DOS DADOS DE EXEMPLO SVM.	76
GRÁFICO 2 – GRÁFICO DOS RESULTADOS DAS ANÁLISES.....	78
GRÁFICO 3 – GRÁFICO SUBGRUPOS.....	83
GRÁFICO 4 – GRÁFICO MSE'S DAS RNA'S CRIADAS.....	92
GRÁFICO 5 – GRÁFICO RNA N° 123	93
GRÁFICO 6 – GRÁFICO RNA N° 174	94

LISTA DE QUADROS

QUADRO 1 - TERMOS PRESENTES NA DEFINIÇÃO DE KDD.....	50
--	----

LISTA DE ABREVIATURAS E SIGLAS

ABAL	ASSOCIAÇÃO BRASILEIRA DO ALUMÍNIO
ALCAN	ALUMÍNIO DO CANADÁ
IAI	<i>INTERNATIONAL ALUMINIUM INSTITUTE</i>
KDD	DESCOBERTA DE CONHECIMENTO EM BASES DE DADOS OU <i>KNOWLEDGE DISCOVERY IN DATABASES</i>
PH	POTENCIAL DE HIDROGÊNIO
PRODERNA	PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE RECURSOS NATURAIS DA AMAZÔNIA
RNA	REDES NEURAIS ARTIFICIAIS
RP	RECONHECIMENTO DE PADRÕES
SVM	<i>SUPPORT VECTOR MACHINE</i>

SUMÁRIO

1	INTRODUÇÃO	15
1.1	JUSTIFICATIVA DA PESQUISA	17
1.2	OBJETIVOS	19
1.2.1	OBJETIVO GERAL.....	19
1.2.2	OBJETIVOS ESPECÍFICOS	19
1.3	ESTRUTURA DO TRABALHO	20
2.	PROCESSO E PROBLEMA.....	21
2.1	O PROCESSO <i>BAYER</i>	21
2.1.1	A BAUXITA.....	24
2.1.2	A QUÍMICA DO PROCESSO <i>BAYER</i>	25
2.1.3	PRINCIPAIS ETAPAS DO PROCESSO <i>BAYER</i>	26
2.1.4	CARACTERÍSTICAS DO PROCESSO PESQUISADO.....	30
2.2	A PROBLEMÁTICA	32
3	TRABALHOS CORRELATOS	39
4	REFERENCIAL TEÓRICO	49
4.1	DESCOBERTA DE CONHECIMENTO EM BASES DE DADOS (<i>KNOWLEDGE DISCOVERY IN DATABASES - KDD</i>).....	49
4.1.1	<i>DATA MINING</i> (<i>DM</i> OU MINERAÇÃO DE DADOS)	52
4.2	PROCESSO DE CLASSIFICAÇÃO.....	54
4.3	PROCESSO DE PREVISÃO	57
4.4	TÉCNICAS APLICADAS NO ESTUDO	60
4.4.1	REDES NEURAS ARTIFICIAIS.....	60
4.4.1.1	TIPOS DE REDES NEURAS	63
4.4.1.2	TIPOS DE APRENDIZAGEM DE UMA REDE NEURAL	65
4.4.1.3	ALGORITMO <i>BACKPROPAGATION</i>	67
4.4.2	<i>SUPPORT VECTOR MACHINE (SVM)</i>	68
4.4.2.1	APRENDIZADO SUPERVISIONADO DA SVM	73
4.4.2.2	UM EXEMPLO DIDÁTICO DE CLASSIFICAÇÃO COM SVM	74
4.4.2.3	CONSIDERAÇÕES SOBRE APLICAÇÃO COMPUTACIONAL DO SVM	76
5	APLICAÇÃO DAS TÉCNICAS DE RP.....	78
5.1	OBTENÇÃO E ESTRUTURAÇÃO DO BANCO DE DADOS.....	78
5.2	RESULTADOS OBTIDOS NA CLASSIFICAÇÃO.....	84
5.2.1	RESULTADOS COM SVM	84

5.2.2 RESULTADOS COM AS RNAs.....	86
5.3 RESULTADOS OBTIDOS NA PREVISÃO.....	90
5.3.1 RESULTADOS COM RNAs.....	90
6 CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS	95
REFERÊNCIAS.....	97
APÊNDICES.....	102

1 INTRODUÇÃO

A humanidade vem buscando ao longo do tempo, formas mais otimizadas/eficientes para executar as atividades do seu dia-a-dia para reduzir esforços, proporcionar conforto, segurança, saúde, lazer, dentre outros. As organizações produtoras desempenham um papel fundamental nesta busca, pois é através de seus produtos e serviços que a humanidade vem desfrutando de melhores condições de vida.

No entanto, realizar uma otimização, independentemente do processo, não é uma tarefa simples, pois são muitas as variáveis qualitativas e quantitativas que podem, ou não, estarem correlacionadas e, para isso, existem inúmeras técnicas que auxiliam o processo de realização da otimização.

Bassanezi (2002) afirma que quando temos uma variável dependendo quantitativamente de outra variável independente, pode-se construir um modelo matemático e analisar esta dependência através das variações destas grandezas. Neste contexto, têm-se os processos com a utilização de mais de uma variável de controle normalmente correlacionadas entre si, responsáveis pelas transformações ao longo das etapas dos processos.

Em processos com sistemas de controle complexos, devido à grande quantidade de variáveis, é possível utilizar técnicas de Reconhecimento de Padrões (RP) como, por exemplo, Redes Neurais Artificiais (RNA), *Support Vector Machine* (SVM), dentre outras, pelo fato de possuírem uma capacidade de modelar sistemas complexos, sejam lineares ou não, e também por possuírem mecanismos de aprendizagem.

De modo geral, nos processos produtivos sempre existe a presença de variáveis de entrada, que são responsáveis por alguma transformação em determinada etapa do processo e que de acordo com a complexidade estas variáveis sofrem mudanças em quantidades e tipos, podendo ser de natureza eletromecânica ou Físico-química.

As variáveis eletromecânicas podem estar relacionadas aos efeitos físicos que, segundo Pahl *et al.* (2005), podem ser descritos quantitativamente por meio de leis físicas que coordenam entre si as grandezas envolvidas, como exemplo, o efeito alavanca através da lei da alavanca, o efeito de dilatação pela lei de dilatação dos corpos. Os desvios nestas variáveis podem gerar perdas de produções imediatas, reduzindo assim a produtividade das instalações.

Em processos produtivos onde a predominância das variáveis é de natureza físico-química, observa-se que os resultados desejados nem sempre são atingidos em função de desvios em apenas uma das variáveis, o que às vezes torna o sistema muito frágil. Este fato se deve principalmente pela correlação entre as variáveis. O impacto negativo no processo produtivo será proporcional ao grau de relevância da variável que apresenta o desvio.

Um dos segmentos industriais que tem como principal característica a utilização de fenômenos físico-químicos para a obtenção de seu produto, é a indústria do alumínio, que segundo Luz e Lins (2005) é o terceiro elemento em abundância na crosta terrestre, ficando atrás do oxigênio e silício.

Para a obtenção do alumínio metálico existe uma etapa muito importante, que é a retirada do óxido de alumínio (Al_2O_3) da rocha de bauxita. O processo mais utilizado industrialmente nesta etapa é o Processo *Bayer*, que consiste em misturar a bauxita moída a uma solução de soda cáustica, com a qual a mesma reage sob pressão e temperatura, dissolvendo a bauxita e formando uma solução de aluminato de sódio. (LUZ e LINS, 2005)

Dentro do complexo processo de extração da alumina destaca-se uma das etapas de fundamental importância para o bom desempenho do processo, denominada de “Lavagem e Filtração da lama vermelha”, que tem como principal finalidade, a retirada de solução de aluminato de sódio contido no rejeito (lama vermelha) e seu reaproveitamento no processo produtivo.

Os filtros rotativos que compõem o processo de lavagem da lama vermelha operam com a utilização de uma grande quantidade de variáveis que combinadas promovem a retirada de soda cáustica desta lama e, em seguida, é transferida para uma bacia de contenção a céu aberto preparado para tal fim.

Para melhor eficiência dos filtros rotativos e redução do teor de soda cáustica na lama vermelha, mantendo-o dentro de um limite tolerável, é necessário que sejam realizadas alterações nos parâmetros de processo, de tal forma que, uma nova combinação das variáveis mantenha a eficiência do mesmo.

Existem ferramentas que empregam métodos exatos de Programação Matemática capazes de nos mostrar de forma quantitativa, a melhor solução para a utilização de variáveis de entrada em um processo produtivo, buscando a maximização de sua eficiência.

Assim, a aplicação do RP neste processo auxiliará a determinar a melhor combinação entre as variáveis, permitindo a otimização da produção, conseqüentemente, aumento da eficiência desta unidade.

A utilização de técnicas tradicionais para otimização ou soluções de problemas pode se tornar inviável pela elevada quantidade de variáveis aplicadas em um único modelo e o elevado tempo de espera para obtenção da resposta. De acordo com Braga *et al.* (2011), em casos mais complexos a utilização de métodos de aprendizagem como as Redes Neurais Artificiais tem obtidos melhores resultados, pela facilidade que as mesmas têm em executar modelagens com um grande número de variáveis não lineares em um espaço de tempo relativamente curto e com um bom grau de precisão.

Este trabalho tem por objetivo aplicar técnicas de RP em dados históricos, visando reconhecer no rejeito do processo de produção de alumina, denominado lama vermelha, os diferentes teores de soda cáustica gerados pelas diferentes entradas, bem como realizar a previsão deste resultado, a fim de antecipar ações que possam minimizar o teor de soda cáustica contida na lama vermelha.

Para a definição do melhor método que se adéqua às necessidades do processo, serão utilizadas as técnicas de Redes Neurais Artificiais e *Support Vector Machine*. Seus resultados serão comparados e aquele que oferecer o maior taxa de acertos em relação aos dados laboratoriais poderá servir de base para a criação do modelo matemático e assim realizar as simulações e verificar a possibilidade de redução do teor cáustico na lama vermelha, reduzindo assim os impactos ambientais.

1.1 JUSTIFICATIVA DA PESQUISA

O alumínio é um metal a cada dia mais utilizado nos segmentos industriais, o que torna seu consumo, cada vez maior, principalmente pela descoberta de novas ligas metálicas, as quais possibilitam a sua aplicação nos mais variados segmentos industriais e atividades em geral, sejam de natureza econômica ou de uso doméstico. Ele é produzido em escala comercial desde o século XVIII e seu mercado vem se expandindo desde então.

Atualmente, o mercado possui 46 países produtores de alumínio, onde o Brasil é o sexto produtor mundial, possuindo a terceira maior reserva do minério

(Bauxita) que está localizada na região amazônica e em reservas que podem ser encontradas no sudeste do Brasil, na região de Poços de Caldas e Goitacazes (MG). A maior parte do alumínio é aplicada nos segmentos de embalagens e transportes, seguidos do setor elétrico, construção civil, bens de consumo, máquinas e equipamentos (ABAL, 2013).

O alumínio para ser produzido necessita de alumina, que é extraída da Bauxita por um processo químico composto por várias etapas de transformações, até atingir o estado de alumina. Nas diversas refinarias espalhadas pelo mundo, o processo é muito semelhante, porém algumas etapas são bem específicas, de acordo com o tipo de Bauxita processada.

Uma das etapas que é comum em praticamente todas as refinarias de Bauxita no mundo é o processo de lavagem do rejeito (denominado de lama vermelha) responsável em retirar o máximo possível da solução de aluminato de sódio (soda cáustica). Este processo é realizado com filtros rotativos que utilizam condensados para a lavagem, uma água proveniente do resfriamento do vapor vindo da etapa de digestão do mesmo processo produtivo, e um sistema de absorção por vácuo.

Além do condensado e o sistema de vácuo, outras variáveis são levadas em consideração pela engenharia de processo para o controle do teor cáustico. O controle destas variáveis é feito através de um sistema automatizado, embasado em gráficos de controle, que mantém os limites máximo e mínimo de cada variável. Tais limites são definidos pela engenharia de processo, porém, quando ocorrem desvios que ultrapassam os limites, os controladores não são mais capazes de efetuar a correção, sendo necessária então a intervenção humana.

O controle de teor cáustico no rejeito do processo de alumina é um dos desafios enfrentados pelas empresas deste ramo, que está diretamente ligado ao controle ambiental da organização. A soda cáustica é um dos principais insumos no processo de produção da Alumina, que é matéria prima na produção do Alumínio.

A necessidade de um controle efetivo do teor cáustico auxilia não somente na redução de custos da organização, pois cada tonelada de soda cáustica que fica retida no rejeito necessita ser repostada no processo, a qual tem um elevado custo de aquisição, como também reduz os impactos ambientais causados pelo produto químico. Com a redução deste teor cáustico na lama vermelha haverá a possibilidade desta ser utilizada na fabricação de cerâmicas e na construção civil, já que poderá fazer parte da composição do concreto.

A presente pesquisa se justifica pelo fato das intervenções humanas serem baseadas no conhecimento adquirido e experiência dos operadores que, muitas vezes, não são padronizadas em todos os turnos de trabalho, fazendo assim, que a tomada de decisão seja pouco eficaz, acarretando perdas de produção e/ou de baixa eficiência na área de lavagem da lama.

Com isso, este estudo servirá de referência para a uniformidade das decisões e que estas ocorram de forma sistemática/automática, com base nos padrões reconhecidos, não havendo a necessidade de intervenção dos operadores.

Com o estudo pode-se prever o resultado do teor cáustico no rejeito, além de ser possível realizar simulações dos resultados esperados na variável de saída reduzindo ao máximo o desperdício de soda no rejeito do processo, possibilitando o uso da lama vermelha em outros processos produtivos.

1.2 OBJETIVOS

Os objetivos deste trabalho podem ser discriminados em geral e específicos, conforme a seguir.

1.2.1 OBJETIVO GERAL

Utilizar técnicas de RP (Redes Neurais Artificiais e *Support Vector Machine*), de forma comparativa, para reconhecer no rejeito do processo de produção de alumina, os diferentes teores de soda cáustica gerados pelas diferentes entradas. Com base no método de RP com melhor desempenho será possível minimizar o teor cáustico contido no rejeito, através da melhor combinação entre variáveis de entrada do processo aumentando, desta forma, a eficiência da lama vermelha.

1.2.2 OBJETIVOS ESPECÍFICOS

- ✓ Identificar as variáveis que interferem no controle do teor cáustico do rejeito do processo;
- ✓ Aplicar as técnicas de RP citadas (RNA e SVM), para identificar qual delas melhor se adéqua e que apresenta o melhor desempenho aos parâmetros que realizam o controle cáustico no rejeito;

- ✓ Aplicar a técnica de previsão com RNA, afim informar aos especialistas de processo os possíveis problemas no sistema e necessidade de ações corretivas;
- ✓ Comparar a atual metodologia que vem sendo utilizado pela indústria pesquisada com o modelo proposto, explicitando os possíveis ganhos reais;
- ✓ Identificar os pontos de melhoria no sistema de lavagem de lama objetivando a redução da concentração cáustica minimizando, desta forma, os impactos ambientais e redução de custos com a reposição de soda cáustica.

1.3 ESTRUTURA DO TRABALHO

Este trabalho está organizado em seis capítulos, além deste capítulo introdutório. O capítulo 2 apresenta a descrição de todas as etapas do processo Bayer, até o produto final, a Alumina, bem como a caracterização do problema em estudo, os efeitos negativos ao meio ambiente, suas características e suas variáveis de controle.

O capítulo 3 apresenta os trabalhos correlatos, abordando trabalhos que também fizeram uso de técnicas de RP. O capítulo 4 apresenta o referencial teórico das técnicas aqui aplicadas, onde são detalhados os seus conceitos.

O capítulo 5 traz como se procedeu a coleta e estruturação dos dados da problemática, a aplicação das técnicas de RP, tanto para a classificação quanto para a previsão de padrões, ao problema estudado e seus resultados são apresentados.

O capítulo 6 aborda conclusões da referida pesquisa, assim como as perspectivas para trabalhos futuros.

2. PROCESSO E PROBLEMA

Neste capítulo aborda-se os conceitos básicos do processo de produção de alumina, denominado Processo *Bayer*, bem como a área específica dentro do processo *Bayer* onde está inserida a problemática deste trabalho.

2.1 O PROCESSO BAYER

Dentre os processos industriais, este estudo enfoca o processo de uma refinaria de produção de alumina, denominado Processo *Bayer*, no qual seu histórico é relevante para o entendimento do estudo em questão, pois a maior parte da alumina produzida neste processo é utilizada para a fabricação de alumínio primário. Por esse motivo faz-se necessária uma abordagem sobre o processo *Hall Héroult*, utilizado para a produção de alumínio. Assim fechando o ciclo do Alumínio e entendendo melhor a geração do rejeito do processo, denominado de lama vermelha.

Segundo o Manual da empresa Alumínio do Canadá, ALCAN (1994), o alumínio é o terceiro elemento mais abundante na crosta terrestre (8,1% em massa) e está presente em combinações muito estáveis na natureza, principalmente com óxidos e silicatos. O primeiro estudo sistemático sobre a existência do alumínio foi do cientista Humphry Davy que em 1808 patenteou o nome desse elemento químico. O alumínio já adotou várias nomenclaturas e processos de obtenção.

Há registros de que em 1825 o norte-americano Charles Martin e o francês Paul-Louis-Toussaint Héroult trabalharam em um processo eletrolítico para a obtenção de alumínio. Estes estudiosos não se conheciam nem sabiam que trabalhavam com a mesma teoria, porém descobriram que o óxido de alumínio se dissolvia em criolita fundida formando uma solução na qual se aplicava uma corrente elétrica para formar a eletrólise, sendo o alumínio depositado no fundo da solução. Logo após a descoberta de Hall e Héroult, houve várias experiências em redução conduzidas pelas fábricas de alumínio (ALCAN, 1994).

Segundo Bliss (1993), o Processo *Bayer* é, atualmente, o único processo economicamente viável para a produção de alumina, e foi desenvolvido pelo químico austríaco Karl Joseph Bayer que o patenteou em 1889. Este estudioso desenvolveu o processo em grande escala de produção de alumina a partir da bauxita. O invento é de grande valor para as refinarias de alumínio, pois necessita do óxido de alumínio

(Al_2O_3) produto final do Processo *Bayer*, para a produção do alumínio metálico produzidos em cubas eletrolíticas ou nos comumente chamados de fornos de redução de alumínio.

No mercado da alumina, o Brasil ocupa a quarta posição mundial, produzindo em 2010 valores superiores a nove milhões de toneladas. Deste total, o estado do Pará foi responsável por mais de 60% da produção nacional (ABAL, 2013).

Algumas propriedades fundamentais para o entendimento do processo de extração da alumina através do ataque da soda cáustica na bauxita são descritas a seguir. Tais propriedades estão relacionadas com solutos e solventes.

a) Solubilidade

Toda solução é constituída de um soluto e um ou vários solventes. Solute denomina-se a parte da solução diluída pelo solvente. Solvente, geralmente, é a substância que está em maior quantidade em uma solução, portanto é considerado como substância que dilui o soluto. Porém, para se obter soluções verdadeiras e para que seja verificada a ocorrência dessa solvência, é fundamental avaliar o grau de afinidade entre as mesmas, partindo do princípio de que se deve analisar antecipadamente as substâncias envolvidas e determinar qual a natureza de cada componente das substâncias envolvidas.

A medida de solubilidade é a relação de quantidade de soluto e de solvente existente na reação estudada. Se a reação em análise apresentar pouca quantidade de soluto em relação ao solvente, constituir-se-á uma solução insaturada. Por outro lado, grandes quantidades de soluto em relação ao solvente produzirão uma solução supersaturada, com crescimento de cristais na solução. Logo, o grau de saturação é identificado pela concentração das substâncias da solução e influencia o meio em que se dá a reação.

b) Efeitos da Temperatura

A temperatura influencia diretamente numa solução em equilíbrio, fazendo com que a reação tente alcançar um novo equilíbrio no sentido de fuga dessa perturbação (MAHAN & MYERS, 2000). A vibração das partículas do soluto e do solvente ocorre devido ao aumento de temperatura estabelecendo o quanto essas moléculas devem interagir através de choques mecânicos entre si, deixando a solução mais dispersa e favorecendo a saturação. Caso contrário, diminuindo a

temperatura de equilíbrio da reação haverá um favorecimento à supersaturação do soluto na solução correspondente, já que a velocidade das partículas menores possibilita melhor adequação facilitando o agrupamento.

c) Efeitos da Concentração

De acordo com o princípio de *Lê Châtelier*, a concentração também exerce efeito direto na reação em equilíbrio, pois, se variar a concentração de um dos componentes da reação à mesma tenderá a ativar esse acréscimo, ou seja, resultará em um aumento da concentração no sentido oposto da reação que fora alterada essa concentração (MAHAN & MYERS, 2000).

Partindo desses princípios, uma solução, ao sofrer uma perturbação em suas partículas, causada por uma variação de temperatura, reage expandindo suas moléculas e tornando o soluto mais diluído pelo solvente que o envolve, aumentando a concentração de solvente e diminuindo a concentração do soluto. É importante salientar que a solubilidade sofre outras influências como efeito do íon comum, pH (potencial de hidrogênio, que indica a acidez, neutralidade ou alcalinidade de uma solução aquosa), mas não é objetivo deste trabalho adentrar no mérito da questão.

d) Solubilidade da Bauxita

A solubilidade da bauxita depende do solvente em que ela esteja dissolvida. No caso do Processo *Bayer*, é a solução de soda cáustica, tendo uma concentração do íon aluminato resultante como produto. Vale ressaltar que, no Processo *Bayer* tal reação química, sofre o efeito da temperatura e da concentração, como fora mencionado anteriormente.

Após a dissolução da bauxita, são explorados os principais minérios de bauxita aluminosa, que são: Gibsita; Boehmita; e Diásporo; tais compostos já dissolvidos em solução de soda cáustica, ($NaOH$), sob determinadas condições de pressão e temperatura, ocorre a solubilidade do tri-hidrato de alumínio ($Al(OH)_3$), durante a dinâmica do processo até a calcinação, processo a ser descrito na explicitação do processo de refinamento da bauxita. ALCAN (1994).

2.1.1 A BAUXITA

A bauxita foi descoberta em 1821 pelo químico francês P. Berthier, na localidade de Lês Baux, no sul da França. Em 1847, Durfrenoy chamou de *Beauxite*, e, em 1861, o metalurgista francês denominou de bauxita. A rocha bauxita (figura 1, a seguir) é composta por uma mistura impura de minerais de alumínio. Esses minerais são conhecidos como oxi-hidróxidos de alumínio, e suas proporções na rocha variam muito entre os depósitos, bem como o tipo e a quantidade de impurezas do minério, tais como: óxido de ferro, argila, sílica, dióxido de titânio, entre outras. A maioria das bauxitas, economicamente aproveitáveis, possui um conteúdo de alumina (Al_2O_3) entre 50 e 55% e o teor mínimo para que ela seja aproveitável é da ordem de 30%.



FIGURA 1 – Minério de Bauxita em sua fase de extração.

FONTE: Hydro (2014).

É importante enfatizar que, na segunda metade do século XIX, quase toda a bauxita era produzida na França e empregada, basicamente, para fins não metalúrgicos. Naquela época, a produção de alumina destinava-se principalmente ao uso como mordente na indústria têxtil. No entanto, com o desenvolvimento do processo *Hall-Hérault* em 1886, a alumina disponível foi empregada de modo crescente usada na produção de alumínio metálico.

Segundo Luz (2003), a produção anual de bauxita é superior a 120 milhões de toneladas, onde cerca de 95% da produção são utilizados na obtenção do alumínio metalúrgico, pelo Processo *Bayer*, seguido do processo *Hall-Hérault*, cujo produto final é o metal, ou seja, o alumínio têm produção estimada em 40 milhões de

toneladas por ano. Uma pequena percentagem de bauxita (5%) não-metalúrgica após calcinação é aplicada na manufatura de abrasivos, refratários, cimento, entre outras.

A alumina utilizada para produtos químicos, não é calcinada, mantendo-se na forma hidratada. De acordo com o *International Aluminium Institute* (IAI, 2013) aproximadamente 40% da produção de bauxita produzida no mundo provêm da Austrália, destacando-se, ainda como principais produtores: Guiana (14%), Jamaica (11%), Brasil (8%), Índia (5%) China (3%). As reservas mundiais de bauxitas estão distribuídas nas regiões tropicais (57%), mediterrânea (33%) e subtropicais (10%).

O Processo *Bayer* extrai a alumina que está contida na bauxita, quando dissolvida em solução de soda cáustica, sob determinadas condições de pressão e temperatura. Porém o processo apresenta uma melhor extração, dependendo da natureza do minério de bauxita, devido suas diferentes características químicas, sendo que a bauxita do tipo Gibsita apresenta maior solubilidade, ou seja, apresenta uma maior quantidade de alumina extraível. A tabela 1 a seguinte, apresenta as condições empregadas na solubilização dos compostos da bauxita em uma refinaria. Se o minério for constituído de uma mistura de dois ou dos três compostos, as condições de digestão são escolhidas considerando-se os componentes menos solúveis. De acordo com Constantino *et al.* (2002), o processo de extração da bauxita rica em Gibsita é o mais econômico.

TABELA 1 - Análise dos Componentes da Bauxita.

Composição da bauxita	Temperatura (k)	[NaOH] (g/l)	[Al ₂ O ₃] final (g/l)
Gibsita	380	260	165
	415	105 - 145	90 - 130
Boehmita	470	150 - 250	120 - 160
	510	105 - 145	90 - 130
Diásporo	535	150 - 250	100 - 150

FONTE: (Constantino *et al.*, 2002).

2.1.2 A QUÍMICA DO PROCESSO BAYER

Segundo a ALCAN (1994), a química do Processo *Bayer* consiste em digerir a bauxita com solução de soda cáustica (*NaOH*), na presença de aluminato de sódio a altas temperaturas (150°C ~ 250°C) e pressão elevada (10kgf ~ 15kgf), para

solubilização dos minerais que contém Al_2O_3 . Nestas condições, a bauxita dissolve-se formando uma solução de aluminato de sódio ($\text{Na}_2\text{O} \cdot \text{Al}_2\text{O}_3$), enquanto as impurezas permanecem na fase sólida conhecida como “lama vermelha”. Outras impurezas consistem na presença de certos minerais nos concentrados de bauxita, mesmo em concentrações pequenas, que ainda afetam a operação do Processo *Bayer* e a qualidade da alumina. Esses minerais se dissolvem em solução cáustica, proporcionando à lixívia, fósforo, vanádio, zinco e matéria orgânica. Adicionalmente, o processo químico, como um todo, é afetado quando ocorre a interação de outras matérias primas requeridas no processo e o efeito acumulativo de compostos na recirculação do licor (soda cáustica reaproveitada do processo). Isso contribui, de forma negativa para a operação do processo e para a qualidade da alumina.

2.1.3 PRINCIPAIS ETAPAS DO PROCESSO *BAYER*

De acordo com ALCAN (1994) o Processo *Bayer* para a produção de alumina está dividido em etapas. A bauxita depois de armazenada em um pátio de estocagem é introduzida no processo inicialmente pela área da moagem onde sofre adição de soda cáustica, que é destinada tanto para realizar a reação química do processo quanto para viabilizar o transporte pelas tubulações. O processo finaliza quando a alumina, extraída da bauxita, sai da etapa de calcinação. O fluxograma da figura 2 representa resumidamente as etapas do processo *Bayer* e na sequencia será exposta uma breve descrição das etapas do processo, sendo o mesmo válido para qualquer outra refinaria que utilize este processo.

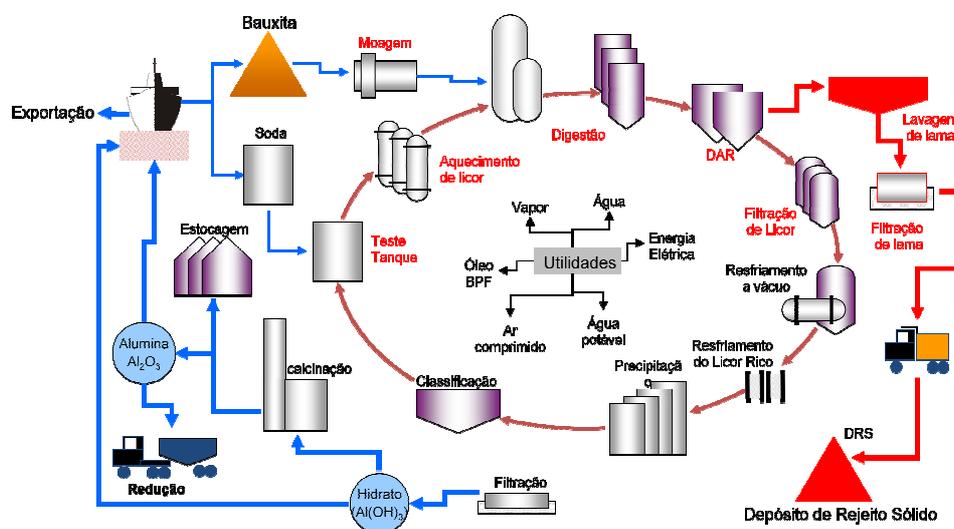


FIGURA 2 – Fluxograma do processo *Bayer*.
 FONTE: Adaptado de IAI (2014)

A bauxita é enviada do pátio de estocagem para o moinho que pode ser do tipo barras e bolas, com auxílio de correias transportadoras que se movimentam sobre roletes. O **processo de moagem** tem como objetivo reduzir a granulometria das partículas, ou seja, reduzir o grão da bauxita para uma melhor ação da soda cáustica que fará a reação química na etapa seguinte denominada Digestão. Nos moinhos é adicionada, juntamente com a bauxita, uma solução de cal hidratada ($Ca(OH_2)$), que tem como principal objetivo promover a diminuição de íons, carbonato e fosfato que estão presentes na bauxita e que afetam a qualidade da alumina, além disso também adiciona-se uma proporção de licor pobre (soda cáustica reutilizada do próprio processo *bayer*, com baixa taxa de concentrado de alumina), este licor pobre irá auxiliar na formação de uma polpa de bauxita que facilitará o transporte pelas tubulações, bem como auxiliará na reação química para a extração da alumina.

Após a moagem, a polpa de bauxita entra na etapa do **processo de digestão**, por meio do qual a rocha é desintegrada pelo ataque de licor pobre a altas temperaturas e pressão. Esta etapa de digestão assemelha-se ao processo digestivo do corpo humano. Com as elevadas temperaturas o licor pobre reage quimicamente na bauxita, fazendo a extração da alumina. Nesta etapa a polpa de bauxita contida nos digestores (equipamento que realiza a digestão) está basicamente dividida em alumina dissolvida (aluminato de sódio solúvel) e polpa de lama, contendo os rejeitos do processo.

Durante este processo de digestão, o rejeito contém sílica reativa proveniente da própria composição da bauxita que necessita ser retirada do processo em uma etapa específica. O resultado da precipitação da sílica reativa a altas temperaturas, que ocorre no decorrer do processo de digestão, é medido no último digestor do referido processo. A sílica reativa é responsável por 90% do consumo de soda cáustica no processo (ESCOBAR, 2004). Exercendo importante papel no controle da qualidade da alumina.

O óxido presente na bauxita em forma de caulinita ($Al_2O_3.SiO_2.2H_2O$), determinada na indústria como sílica reativa, também é dissolvido nas condições da digestão formando silicato de sódio. Após a dissolução e a supersaturação da solução de processo com o silicato, ele precipita-se na forma de um óxido triplo de sódio, alumínio e sílica, conhecido na indústria como produto de dessilicação. O processo de extração da alumina contida na bauxita, um tempo estimado de 10

minutos, porém necessita de uma permanência de aproximadamente 60 minutos neste processo, dependendo do teor de sílica reativa, tal tempo é necessário para a precipitação sílica não reativa.

No processo de digestão é necessário elevar à temperatura a cima de 140°C. Em seguida há uma necessidade de fazer o **resfriamento desta polpa**, com isto, na saída do último digestor há tanques de expansão que recebem a polpa com temperatura e pressão elevada, e utilizando um simples processo, realiza o resfriamento e deixa a mesma com temperatura aproximada de 108°C e pressão atmosférica (ALCAN, 1993). Este resfriamento serve para evitar a autoprecipitação na etapa de decantação seguinte, além disso, este processo de resfriamento proporciona a liberação de vapor que posteriormente condensa voltando ao estado líquido, tal condensado irá alimentar a área de lavagem de lama, tornado assim um processo de circuito fechado, minimizando custos no consumo de água e auxiliando na sustentabilidade ambiental do processo.

Na etapa seguinte, a polpa resfriada a uma temperatura de aproximadamente 108°C proveniente da Digestão, passa por tanques de **decantação**, separando o resíduo sólido rico em óxido de ferro (lama vermelha) da solução de aluminato de sódio, ($Na[Al(OH)_4]$), ocorrendo uma **clarificação** da solução de transbordo. No fundo dos decantares fica a lama vermelha. Tal processo é de simples entendimento, pois tais decantadores são tanques com fundo cônico, isso facilita a lama vermelha a ser depositada no fundo pela ação da gravidade e com auxílio da adição de um agente sintético denominado floculante. Após a lama vermelha estar depositada no fundo dos decantados ela passa por um processo de **lavagem e filtração em filtros à vácuo** de tambor rotativo para a recuperação da soda cáustica contida na lama. A lama vermelha após passar pelo processo de lavagem ela é acumulada em depósitos apropriados para tal fim. Tal processo será melhor abordado na seção 2.2 onde trata-se a problemática deste estudo.

A lama vermelha depositada no fundo dos decantadores segue para o processo lavagem, porém a parte líquida que contém o licor rico (soda cáustica com aluminato de sódio, Alumina) segue para a etapa de **filtração do licor rico** que tem por objetivo eliminar quaisquer indícios de sólidos ainda suspensos no licor rico provindo dos decantadores, através de filtração. O filtrado é, então, resfriado e o ($Al(OH)_3$) é precipitado pela adição de partículas (germes de cristalização) de hidróxido de alumínio.

Após o licor rico sofrer o processo de filtração, ele continua com a temperatura elevada, em aproximadamente 100°C, que até neste momento tal temperatura tem a finalidade de manter a alumina em suspensão no licor rico. Após a filtração, o licor rico necessita ser resfriado para iniciar o processo de separação da alumina da soda cáustica, formando os cristais de alumina hidratada denominado de Hidrato. Nesta etapa de **resfriamento a vácuo** o licor rico é resfriado através de tanques de expansão reduzindo sua temperatura, ocasionando a redução do ponto de ebulição devido ao decaimento de pressão, executado por bombas de vácuo. Nesta etapa é observada uma redução da temperatura do licor, além de uma elevação da supersaturação de $Al(OH)_3$ formando o hidrato.

Conforme ALCAN (1994), na etapa de **precipitação**, a temperatura do licor já se encontra reduzida, resultante do processo de resfriamento a vácuo e, portanto com grau de supersaturação elevado. A partícula de hidrato ($Al_2O_3 \cdot 3H_2O$) se forma com a sua superfície granulométrica particular, de acordo com a especificação do cliente, a partir do controle de temperatura e da relação da concentração de alumina em função da concentração cáustica em tanques precipitadores que tem a função de precipitar os cristais de hidrato formando uma polpa de hidrato.

Em seguida ocorre o processo de **classificação do hidrato** que tem por finalidade, separar a semente de hidrato e enviá-la para a filtração. As sementes, fina e grossa, recirculam na própria precipitação. Nesta etapa ocorre a separação física do hidrato e da soda cáustica, que passa a se chamar de licor pobre, pois possui uma baixa concentração de alumina, este licor pobre será destinado ao processo de digestão da bauxita, fechando o ciclo do licor. Essa etapa de clarificação do hidrato é composta de tanques classificadores (primários e secundários) e espessadores.

O hidrato mesmo após ser fisicamente separado do licor, ainda continua com elevado teor de umidade. Nesta etapa de **filtração do hidrato** ocorre a retirada do máximo de umidade contida na polpa de hidrato, lavando e secando repetidas vezes com a finalidade de remover os resíduos de soda cáustica que ainda persistem na partícula. O controle da umidade é importante para a etapa de **calcinação** objetivando redução de custo com combustível e partículas mais resistentes à quebra (ALCAN, 1994).

Calcinar, diz respeito a tudo aquilo que é aquecido pelo fogo até secar. No caso da calcinação do hidrato significa aquecê-lo até ocorrer a retirada da molécula

de água contida em sua estrutura química. O hidrato ($Al_2O_3 \cdot 3H_2O$) é calcinado para produzir alumina (Al_2O_3). Portanto, a calcinação é a etapa final do processo para a eliminação total de água em fornos calcinadores a uma temperatura superior à $900^\circ C$, ou seja, é o momento da retirada de três moléculas de água do óxido de alumínio, formando o produto final do processo *Bayer*, a figura 3 mostra a alumina calcinada.



FIGURA 3 – Alumina Calcinada.

FONTE: Hydro (2014).

2.1.4 CARACTERÍSTICAS DO PROCESSO PESQUISADO

A lama vermelha, antes de ser descartada na bacia de rejeito sólido, passa por um processo de lavagem com condensado, a fim de reduzir ao máximo a concentração cáustica contida na mesma, e reutilizar este produto cáustico no processo com o objetivo reduzir os custos com a reposição de soda cáustica. Para a lavagem da lama vermelha, além do condensado, outra série de variáveis são combinadas para o melhor desempenho da atividade, tais variáveis serão descritas a seguir.

- a) **Condensado ou Spray:** A água que realiza a lavagem da lama vermelha é reaproveitada do processo de resfriamento da polpa, que gera vapor e após resfriar transforma-se em condensado. Este condensado é bombeado por um sistema de tubulações e é lançado sobre o filtro rotativo que está envolto por

- uma camada de lama vermelha. Tal condensado tem a função de penetrar na lama vermelha, lavando a mesma e reduzindo seu teor cáustico.
- b) **Vácuo:** O sistema de vácuo tem como principal objetivo formar uma camada da lama vermelha na parte externa do filtro rotativo. Como o vácuo é uma pressão negativa, ela succiona o condensado que lavou a lama e transfere para um tanque de armazenagem. O sistema de vácuo atua na parte interna do filtro, facilitando sua atuação e formação da camada de lama.
 - c) **Rotação do Filtro:** Os filtros são cilíndricos do tipo rotativo, e possuem uma rotação variável. Tal rotação é de fundamental importância para a eficiência da lavagem, pois uma rotação elevada permite que a lama passe rapidamente pelos jatos de condensado, impedindo uma boa lavagem. Entretanto uma rotação baixa faz com que a camada de lama sobre o filtro fique muito espessa devido a atuação do vácuo, com isso o condensado não penetra em toda a lama, fazendo apenas uma lavagem superficial, o que também gera uma lavagem ineficiente.
 - d) **Densidade da Polpa de lama:** A polpa da lama possui uma densidade que pode ser manipulada por intermédio de adição de condensado durante o bombeamento do processo de decantação para o processo de lavagem de lama. Esta densidade é importante para o processo, pois uma densidade baixa significa que há mais líquido que sólido, e conseqüentemente uma maior porcentagem de soda cáustica, o que necessitará de muito mais condensado para lavar, e o contrário também é prejudicial, pois a densidade baixa, informa que há muito sólido na polpa, o que facilitará a formação de uma camada espessa de lama vermelha sobre o filtro, que culminará em uma lavagem ineficiente.
 - e) **Nível da Bacia de Lama:** O filtro rotativo fica parcialmente submerso em uma bacia que contém a polpa de lama. Esta bacia que é constantemente alimentada por polpa de lama, necessita ter um controle efetivo de seu nível, pois níveis muito baixos impedem a formação de uma camada de lama adequada na superfície do filtro e é um indicativo de baixa produtividade do sistema, já um nível de bacia muito elevado favorece a formação de uma camada espessa de lama, o que prejudicará sua lavagem.
 - f) **Teor Total de Soda (TTS):** Esta variável refere-se à quantidade de soda cáustica que a polpa de lama vermelha contém antes de entrar no processo

de lavagem. Este valor não controlado na etapa da lavagem e sim na etapa anterior de Decantação da lama. Porém esta variável torna-se importante na medida em que o sistema de lavagem possui limitações físicas e de projeto, portanto quanto menor o TTS na polpa da lama auxiliará em uma melhor eficiência do sistema de retirada da soda cáustica na lama vermelha.

De posse dessas informações, os especialistas da área em estudo puderam inferir que este conjunto de variáveis combinadas de tal forma resultará no bom desempenho, ou não, do processo de lavagem da lama vermelha. Na figura 4 seguinte, ilustra-se o filtro rotativo com as variáveis que o envolve.



FIGURA 4 – Processo de lavagem da lama vermelha.

FONTE – Autor (2014)

A determinação da concentração cáustica no rejeito de lama vermelha do processo *Bayer* é realizada por técnica de potenciometria que detecta o quanto de soda cáustica reside na lama vermelha, para isso uma parcela da lama é filtrada no laboratório da empresa onde é submetida à análise para se identifica a quantidade de soda cáustica que esta amostra possui.

2.2 A PROBLEMÁTICA

No segmento da produção da alumina, o Brasil ocupa a quarta posição mundial. No ano de 2010, o Brasil teve uma produção total de alumina de cerca de 9.433.400 toneladas. (ABAL, 2013). Porém nem toda a bauxita processada gera

Alumina, existe um grande percentual de rejeito do processo inaproveitável economicamente, esta quantidade depende da qualidade da bauxita, que influencia diretamente na quantidade de lama vermelha gerada.

Há uma relação de produtividade existente na produção da alumina, sendo que cada tonelada de alumina produzida, necessita de duas toneladas de bauxita. Vale salientar que estes valores variam de acordo com a especificação de cada tipo de bauxita.

Assim, cada tonelada de bauxita contém um percentual de alumina aproveitável de 50%, sendo que o restante é composto por silicatos, ferro, manganês, potássio e outros metais não economicamente exploráveis (a lama vermelha).

O fato é que a quantidade de lama vermelha gerada anualmente assume enorme proporção, da ordem de milhões de toneladas, representando um sério passivo ambiental, pois, junto a ela, perde-se uma quantidade considerável de soda cáustica, o que também contribui significativamente nos custos de reposição desse insumo, sendo um dos principais insumos com maior custo na produção da alumina.

Em algumas plantas de refino de Alumina, a área desmatada para acondicionar a lama vermelha gerada do rejeito do processo já ocupa uma área maior que a própria área onde a refinaria está instalada, conforme mostram as figuras 5 e 6, e com tendência de crescimento. A produção mundial do alumínio também tende a crescer devido aos novos produtos industrializados que estão cada vez mais utilizando o alumínio em sua composição, e quanto mais alumínio produzido, maior será a produção de alumina que, por sua vez, irá gerar uma elevada quantidade de rejeito, que precisará cada vez mais de espaço para ser acondicionada.



FIGURA 5 – Área de rejeito de lama vermelha / empresa no Estado do Pará- Brasil
FONTE – Imagem de satélite



FIGURA 6 – Área de rejeito de lama vermelha / empresa no Estado do Maranhão- Brasil
FONTE – Imagem de satélite.

Estas figuras 5 e 6 apresentam as imagens de satélite de duas empresas refinadoras de bauxita, onde se observa a área desmatada que é destinada ao acondicionamento da lama vermelha, que já toma as mesmas dimensões da área construída da refinaria. Na figura 7 tem-se a imagem de uma aérea de uma bacia de rejeito de lama vermelha, na qual se pode perceber a extensão da área desmatada para tal finalidade.



FIGURA 7 – Área de rejeito de lama vermelha.
FONTE: Souza (2010).

As bacias de contenção destinadas ao recebimento da lama vermelha são localizadas geralmente próximas à refinaria, a fim de reduzir os custos com o transporte. Estas bacias são planejadas de forma que a lama vermelha não tenha

contato direto com o solo, ou seja, há um revestimento no fundo da bacia com um material de polipropileno, com objetivo de evitar contaminação ambiental, pois como a lama fica exposta ao meio ambiente, em períodos chuvosos, a água da chuva sofreria contaminação e poderia contaminar o lençol freático.

Antes de a lama vermelha ser depositada na bacia de contenção, ela passa por um processo de lavagem, com o objetivo de retirar ao máximo possível os resíduos de soda cáustica, oriunda do processo *Bayer*; a figura 8 ilustra a saída da lama vermelha após o processo de lavagem. Porém mesmo após este processo de lavagem, a lama vermelha ainda contém resíduos de soda cáustica, cuja quantidade, dependendo do processo, atinge em média a faixa de 15 g/l. Tal valor é considerado elevado, pois a soda cáustica é um contaminante. Um resultado ideal seria um valor nulo de soda cáustica na lama, porém tal valor é inviável para o processo devido seu alto custo operacional. Dependendo dos distúrbios do processo, este valor poderá chegar em 40 g/l, que além de aumentar a contaminação da lama vermelha, eleva os custos de produção, tendo-se em vista que esta soda cáustica necessitará ser reposta no processo.

Para uma melhor estabilidade do processo, um valor aceitável, que foi determinado pela engenharia de processo, a soda cáustica contida na lama vermelha deverá ser no máximo de 12g/l.



FIGURA 8 – Saída de lama vermelha após lavagem.

FONTE – Material treinamento refinaria de bauxita.

A quantidade ou volume de lama vermelha descartada é proporcional à quantidade de Alumina produzida que, conseqüentemente, irá necessitar de uma maior área desmatada para criação de novos depósitos de rejeito.

Este desmatamento poderá ser amenizado, caso esta lama vermelha seja reaproveitada em outro processo de fabricação, porém devido ao elevado volume de soda cáustica contida na lama e, também, a sua instabilidade, torna-se inviável tal processo. Em 2005, uma das empresas refinadoras de Alumina solicitou um estudo de viabilidade do reaproveitamento da lama vermelha, sendo que os primeiros resultados foram desanimadores.

Outras empresas vêm estudando o uso desta lama vermelha em escala comercial. Mercury *et al.* (2001) realizaram um estudo preliminar sobre o reaproveitamento da lama vermelha na mistura com argila, para produção de cerâmica. Os resultados deste trabalho foram satisfatórios, sinalizando o reaproveitamento da lama vermelha.

Outros estudos também estão sendo realizados por instituições de ensino em parceria com as empresas. Na Universidade Federal do Pará, alunos e professores dos Cursos de Engenharia Civil e do PRODERNA (Programa de Pós Graduação em Engenharia de Recursos Naturais das Amazônia), vêm estudando o tema, como exemplos, Rossi *et al* (2008), Souza (2010) e Macêdo *et al.* (2011), onde por meio do aproveitamento da lama vermelha na mistura para produção de agregados do concreto estão obtendo resultados satisfatórios.

Com uma mistura de 40% de argila e 60% de lama vermelha, Macêdo *et al.* (2011) verificaram a possibilidade de fabricação de tijolos (blocos), como mostrado nas figuras 9 e 10.



FIGURA 9 – Tijolos (Blocos) produzidos com a lama vermelha.

FONTE: Macêdo *et al.* (2011).



FIGURA 10 – Tijolos (Blocos) produzidos com a lama vermelha.
FONTE: Macêdo *et al.* (2011).

Com objetivo de ratificar a possibilidade de uso comercial deste tipo de tijolo, Souza (2010) mostra que uma casa chegou a ser construída com os tijolos, conforme ilustra a figura 11.



FIGURA 11 – Casa construída com resíduo do processo Bayer.
FONTE: Souza (2010).

Assim, objetivando um melhor controle e previsão do teor cáustico na lama vermelha e com a possibilidade de manipulação das variáveis de entrada, pode-se reduzir ao máximo este valor, fazendo com que a lama vermelha possa ser aproveitada economicamente, reduzindo os impactos ambientais. Pretende-se implantar por meio desse estudo, uma nova metodologia de monitoramento e previsão de resultados, a partir da concepção de um modelo matemático baseado em análise do comportamento das variáveis de controle.

É importante ressaltar que o emprego de tecnologias nos processos é fundamental, pois permitem melhorar o mesmo garantindo o aperfeiçoamento da produção (SLACK e LEWIS, 2009). Os mesmos autores ressaltam a importância de se usar estratégias que proporcionem vantagens no emprego de tecnologias de forma a direcioná-las aos objetivos da empresa.

Da mesma forma, entende-se que um confiável sistema de monitoramento e obtenção de dados e informações, deve ser utilizado por seus gestores para controle eficiente de variáveis, a fim de se obter como retorno, ações eficazes que reduzam as perdas de insumos importantes e contribuam com a redução nos custos de produção.

No campo da área de Pesquisa Operacional, ações de monitoração conclusivas e eficientes ao bom controle de um processo tem sido o foco de conhecimentos baseados em modelagem e simulação, tema dessa pesquisa.

3 TRABALHOS CORRELATOS

Este capítulo relata alguns trabalhos correlacionados com o tema desta pesquisa e traz um resumo de como autores trabalham com as técnicas de RP aqui apresentadas em diferentes problemáticas e seus resultados.

Campos (2001) utilizou a técnica de RP como ferramenta para reconhecimento de faces baseado em características específicas do rosto em análise de imagens gravadas, sendo assim possível reconhecer um rosto após leitura destas características.

Em sua tese, Loh (2001) desenvolveu um método para descoberta do conhecimento em textos (KDT – do inglês *Knowledge Discovery in Texts*) baseado em RP, identificando características em alto nível nos textos na forma de conceitos, analisando termos codificados a um determinado padrão.

Estudos para desenvolver algoritmos para sistemas de RP foi realizado por Castro e Prado (2002), com ênfase em técnicas de agrupamento como “Similaridade Máxima”, “MaxiMin-Distância”, “K-Means” e “ISODATA”. O algoritmo desenvolvido mostrou-se satisfatório para os objetivos do estudo.

Ballini *et al.* (2003) apresentam um estudo utilizando um modelo de rede neural nebulosa para previsão de vazões sazonais. O modelo é baseado em um método de aprendizado construtivo onde grupos de neurônios competem quando a rede recebe uma nova entrada. Os resultados mostram que a rede neural nebulosa forneceu um melhor desempenho para prever um passo à frente, com erros significativamente menores que as outras abordagens.

Constant *et al.* (2004) fizeram uso das redes neurais artificiais (RNAs) para realizar um estudo em um processo de compósitos poliméricos (plásticos reforçados), precisamente no processo de *filament winding* (filamento contínuo ou enrolamento filamentar), para predição do comportamento térmico de tubos compósitos durante a etapa de cura. Os resultados obtidos com as redes neurais foram comparados com os resultados experimentais reportados na literatura, mostrando valores satisfatórios para tal experimento.

Rezende (2005) exemplifica a utilização de RNA em alguns estudos de casos onde as mesmas foram aplicadas e demonstraram bons resultados, como na modelagem para controle de um processo de flotação de minério de ferro da Cia Vale de Rio Doce, onde os resultados obtidos foram animadores. O modelo aplicado em uma das plantas foi colocado em operação *online* com os controladores

convencionais para testar sua capacidade de prever o teor de um dos elementos (sílica) no produto concentrado. Foi realizada a comparação entre os valores previstos pela rede com os valores obtidos através da análise em laboratório. Os resultados da aplicação do modelo demonstram a possibilidade de dispensar a necessidade de uso do laboratório, pelo grande grau de aderência do modelo.

Coelho *et al.* (2005) desenvolveram um modelo de previsão para a resistência mecânica do coque utilizado na indústria metalúrgica, com auxílio da regressão múltipla, em escala-piloto. O modelo de previsão obteve um aumento de 8,07 pontos percentuais em taxa de acerto em comparação com o modelo atual. Portanto o novo modelo desenvolvido mostrou-se adequado e pode ser usado como ferramenta para auxiliar na tomada de decisão dos gestores.

Santos *et al.* (2005) desenvolveram um sistema para predição da soroprevalência da Hepatite A utilizando modelos de regressão de logística e Redes neurais artificiais. O desempenho de tais modelos é medido através da taxa de classificação incorreta em uma amostra do município de Duque de Caxias, Rio de Janeiro, que possui elevada prevalência da doença. A proporção de concordância total, ou seja, a proporção de indivíduos no conjunto de teste que foram classificados como soropositivo (soronegativo), sendo realmente portadores da doença (não portador da doença), para a RNA foi igual a 88% e a regressão logística apresentou uma proporção de concordância total igual a 83%. De acordo com os resultados apresentados observa-se que a RNA apresenta boa sensibilidade, ou seja, ela consegue classificar eficientemente os soropositivos para hepatite A. Em relação à especificidade, ambos os modelos apresentaram uma especificidade bastante elevada.

Lima (2006) realizou um estudo para estimar a concentração de particulado em suspensão em diferentes bairros da cidade de Ipatinga-MG, através da utilização de modelagem com rede neural artificial (RNA). O modelo foi capaz de fazer previsões de concentração de alguns poluentes devido ao seu treinamento utilizando as variáveis de entrada como velocidade dos ventos, chuvas e outros. A variável de saída foi medida pelos resultados obtidos das estações meteorológicas. A validação do modelo foi realizada pela comparação entre os resultados reais e os previstos pela RNA, verificando que é possível a utilização deste tipo de modelo para a previsão de concentração de particulado em suspensão.

Barreiros *et al.* (2006) realizaram um trabalho objetivando investigar, de forma qualitativa as oscilações indesejáveis de baixa frequência de um sistemas elétricos de potência das linhas de transmissão utilizando os chamados Estabilizadores de Sistemas de Potência (ESPs), em que regiões de operação do sistema esta última situação ocorre. O conjunto de parâmetros de um ESP a ganhos programados é utilizado para treinar uma RNA, através do conhecido algoritmo de retropropagação do erro. Devido à capacidade da RNA de fazer o mapeamento de qualquer função não-linear com a precisão desejada, é possível observar as variações que ocorrem nos parâmetros do controlador com a mudança no ponto de operação do sistema de potência. Testes realizados em um sistema do tipo máquina síncrona - barramento infinito demonstram que o ESP Neural proposto apresenta um desempenho muito bom, confirmando a principal característica da RNA treinada.

Santos e Senger (2006) apresentam um algoritmo baseado em Redes Neurais Artificiais (RNA) para a implementação de uma proteção de distância de linhas de transmissão de energia, este algoritmo não necessita de nenhum ajuste na topologia ou nos parâmetros da RNA (retreinamento), quando aplicado à diferentes sistemas elétricos. Esta característica diferencia a solução proposta, das demais baseadas em RNA, que são sempre projetadas para operarem numa única e exclusiva linha de transmissão (geometria de torre, condutores, comprimento e nível de tensão pré-determinados), inviabilizando sua implementação num relé comercial. Observa-se que as RNA avaliadas responderam, com a resistência ou a reatância, com erros menores do que 5%. Na maioria dos casos, erros bem inferiores a 3%.

Dominguete *et al.* (2006) apresentaram um novo modelo para a estimativa do teor de fósforo para o Convertedor LD da V&M do Brasil. O modelo obtido, em trabalho anterior, por meio de regressão múltipla é avaliado perante uma nova regressão, desta vez, não-linear, representada por uma RNA. Ao final do trabalho pode-se concluir que a modelagem de processos utilizando Redes Neurais Artificiais é capaz de proporcionar melhores resultados em comparação as técnicas de regressão lineares. Foi possível projetar, com o uso de Redes Neurais, modelos capazes de estimar com relativa qualidade o teor de fósforo no aço a partir de variáveis de processo relativas à operação do convertedor LD da V&M do Brasil.

Guarnieri (2006) aborda o desenvolvimento e avaliação de uma metodologia de previsão de irradiação solar incidente, voltada ao setor de geração e distribuição de energia, constituindo assim, uma pesquisa de meteorologia aplicada. O trabalho

desenvolvido consiste no refinamento da previsão numérica de tempo produzida pelo modelo Eta/CPTEC, visando reduzir as incertezas associadas à previsão de radiação solar deste modelo. Foram empregadas como preditores em modelos de Redes Neurais Artificiais (RNAs) e Regressões Lineares Múltiplas (RLMs). Ambos os modelos demonstraram resultados satisfatório na previsão de radiações solares do modelo ETA.

Steiner *et al.* (2007) analisaram um conjunto de dados de crédito, e usando a técnica de extração de regras NeuroRule e o software WEKA para a extração de regras a partir de uma RNA treinada com objetivo de apresentar ferramentas que possam auxiliar a identificar e prever quais clientes serão adimplentes (ou não) em relação ao crédito em instituições bancárias. Os resultados foram considerados bastante satisfatórios alcançando mais de 80% de acurácia quanto à concessão (ou não) de crédito bancário em todas as simulações.

Baptistella *et al.* (2007) objetivando a estimação de valores venais de imóveis na cidade de Guarapuava - PR, utilizaram as técnicas de redes neurais artificiais fazendo uso de dados do cadastro imobiliário fornecidos pelo setor de Planejamento da Prefeitura Municipal. As Redes Neurais Artificiais desenvolvidas foram do tipo *feedforward*, utilizando o algoritmo de treinamento Levenberg-Marquardt. Os resultados obtidos através das RNAs, quando da separação por setores, foram melhores em relação ao modelo contendo todos os dados, ou seja, quando setorizou-se a amostra, tanto os erros de treinamento quanto de testes, foram significativamente melhores.

Mota (2007) apresenta uma proposta para a avaliação de imóveis urbanos (apartamentos, casas e terrenos) através de técnicas de Análise Multivariada e Redes Neurais Artificiais. O trabalho consiste em: efetuar a padronização dos dados, separá-los em grupos por meio de um Mapa Auto-organizável de Kohonen, aplicar a Análise de Componentes Principais, com 98% de explicação da variabilidade dos dados, para evitar o problema clássico da multicolinearidade e, só então, aplicar em cada grupo uma RNA de Múltiplas Camadas, para determinar os preços de venda dos imóveis e, posteriormente, comparar os resultados obtidos com os encontrados nos referidos trabalhos. Nesta comparação foi possível observar que as Redes Neurais mostraram um melhor desempenho na predição de preços dos imóveis.

Fontes e Mendes (2008) desenvolveram, com base em uma estratégia de controle preditivo não linear aplicada a um reator tanque agitado de polimerização

em lama para a produção de polietileno de alta densidade, sendo o desempenho do controle analisado no sentido de se obter o peso molecular médio numérico, composição de comonomero e temperatura desejados, um algoritmo de controle baseado em redes neurais com topologia “*feedforward*”. A implementação do algoritmo contempla um problema de otimização. Os resultados de simulação mostram o bom desempenho do algoritmo baseado em um modelo neural treinado para a predição da variável de saída apenas um instante de tempo a frente e, ao mesmo tempo, atestam a dificuldade de controlar diretamente as propriedades do polímero com a ocorrência de tempo morto na medição.

Steiner *et al.* (2008) utilizaram uma metodologia, composta por técnicas de Análise Multivariada, para a construção de um modelo estatístico de regressão linear múltipla para avaliação de imóveis em função de suas características. É ajustado um modelo de regressão linear múltipla para cada grupo de imóveis homogêneos dentro de cada classe, a metodologia foi aplicada a um conjunto de 119 imóveis, referentes a 44 apartamentos, 51 casas e 24 terrenos, da cidade de Campo Mourão - PR. O modelo para cada grupo homogêneo, dentro de cada classe de imóveis avaliados, apresentou um ajuste adequado aos dados e uma capacidade preditiva bastante satisfatória. A metodologia multivariada apresentada para predições dos preços de imóveis se mostrou viável e altamente apropriada, atingindo-se resultados com níveis de precisão bastante satisfatórios. Desta forma, a mesma poderá servir de apoio tanto a gerentes imobiliários na definição de preços de imóveis, como para pessoas físicas e/ou jurídicas que desejam conhecer mais realisticamente o seu patrimônio.

Lima *et al.* (2008) apresentaram um método para a previsão da distribuição espacial de novos casos de dengue nos estados de Alagoas e Paraíba (nordeste do Brasil) com auxílio da regressão múltipla, a partir da previsão da precipitação obtida por modelos atmosféricos regionais. Foram encontrados resultados satisfatórios para ambas as localidades, onde se observou uma boa relação entre a precipitação prevista e os casos de dengue observados e previstos. Verificou-se também, que a previsão da precipitação utilizando o modelo regional aproximou-se bastante dos valores da precipitação observada nestas localidades, evidenciando que a previsão da distribuição espacial da dengue utilizando-se este método é bastante satisfatória.

O RP por processamento de imagens digitais com auxílio pelas Redes Neurais Artificiais (RNA) para identificar algumas variedades de soja por meio de sua

forma e do tamanho das sementes, foi desenvolvido por Tourian e Padilha (2008). O treinamento e testes da RNA mostraram-se eficientes ao estudo proposto.

Scarpari e Beauclair (2009), com base na cultura da cana-de-açúcar, observou que esta é submetida a diferentes condições ambientais e de manejo durante o seu desenvolvimento, o que afeta diretamente a maturação. E com a necessidade de se quantificar as respostas da cultura aos diferentes estímulos para fins de planejamento. Objetivando suprir estimativas de rendimento ao longo da safra, visando à caracterização das alternativas de manejo, aumentando a eficácia das decisões gerenciais e estratégicas. Desenvolveram-se modelos empíricos capazes de obter estimativas de ATR - Açúcar Total Recuperável ao longo da safra, utilizando dados referentes aos fatores de produção. Foram analisados os resultados de colheita realizados no município de Piracicaba – SP, entre os anos de 1998 e 2003 considerando a maturação, idade do canavial, solos, variedades, florescimento e manejo, utilizando-se de modelos estatísticos de estimativa da maturação para estimar a safra 2004. Todos os modelos de previsão para cana-soca mostraram-se significativos, sendo uma ótima ferramenta de auxílio para o planejamento agrícola otimizado.

Lorenzi (2009) utilizou as técnicas de Regressão Múltipla e Redes Neurais Artificiais para analisar a estrutura do concreto armado utilizando ensaios de velocidade de propagação de pulso ultrasônico, que permite realizar um acompanhamento das características do material ao longo de sua vida útil. Com os dados obtidos pode-se averiguar a uniformidade do concreto, controlar sua qualidade e acompanhar sua deterioração ao longo do tempo. Pode-se observar que os resultados obtidos com a RNA demonstraram-se bastante satisfatório em comparação aos da regressão múltipla e aos métodos tradicionais utilizados. O método estatístico apresentou um coeficiente de determinação (R^2) de 0,40 enquanto que pela RNA apresentou um coeficiente na ordem de 0,90. Os bons resultados da RNA indicam seu elevado potencial para produzir estimativas de resistência do concreto.

Palma e Portugal (2009) utilizaram Regressão Múltipla e Redes Neurais Artificiais para identificar empiricamente os motivos das variações da inflação no Brasil, no período pós-metas de inflação, bem como a previsão da mesma em função das diversas variações das variáveis de entrada. Os resultados obtidos neste trabalho permitem afirmar que a maior influência sobre as expectativas inflacionárias

no período como um todo foi da volatilidade cambial, seguida pela variação no preço das commodities, pela defasagem de ordem 1 das expectativas, pela variação cambial e pela meta. Em menor magnitude, afetam as expectativas o resultado primário do governo, a defasagem de ordem 2 e a taxa Selic.

Em seus estudos, Santos *et al.* (2009) utilizaram informações de dados de sensores remotos orbitais (óticos e de radares) na validação de padrões de uso e cobertura do solo na planície costeira amapaense para fins de mapeamento e reconhecimento da dinâmica natural e antrópica. A análise foi realizada aplicando-se dados de satélite digitalmente processados que permitiu distinguir oito padrões. Tal método torna-se importante para o mapeamento dos ambientes costeiros e fundamental para o reconhecimento da dinâmica na região.

Oliveira *et al.* (2010) aplicaram as técnicas de Redes Neurais Artificiais na tarefa de previsão da produção de álcool no Brasil no ano de 2006, a partir de dados de produção dos anos anteriores. A previsão de produção de álcool favorece uma tomada de decisões no sentido de planejar o volume de estoque para o consumo ao longo dos períodos subsequentes. Os resultados obtidos por meio das Redes Neurais foram comparados com os obtidos utilizando técnicas de séries temporais, foram mais satisfatórios.

Gomes *et al.* (2010) apresentam uma análise dos dados pluviométricos e fluviométricos da área a montante do reservatório de Três Marias, no rio São Francisco, bem como a calibração de um modelo baseado na técnica de Regressão Múltipla e Redes Neurais Artificiais para a previsão de vazões naturais afluentes. Os resultados obtidos mostraram que o modelo calibrado com a técnica de RNA teve um desempenho superior ao modelo estocástico, tanto a nível diário quanto semanal.

Beltrami *et al.* (2010) enfatizam que determinar preços teóricos de contratos de opções, mediante modelos matemáticos, permite ao investidor analisar se os preços estabelecidos pelo mercado estão superestimados ou subestimados. Neste contexto, os autores realizaram um trabalho objetivando estabelecer modelos de precificação de opções baseados nas técnicas do *Support Vector Regression* (SVR) e Redes Neurais Artificiais (RNA), comparando seus desempenhos e avaliar suas viabilidades de aplicação no mercado brasileiro. Nessa pesquisa foram utilizados dados referentes às opções de compra americanas sobre ações da Petrobras PN negociadas na BOVESPA, no período de novembro de 2008 a maio de 2009. Os

resultados encontrados mostraram, a partir das comparações realizadas com os valores de mercado, que o SVR foi capaz de determinar valores mais próximos aos de mercado e captar melhor as oscilações do mesmo, obtendo desempenho superior ao da RN na precificação de opções.

Huang *et al.* (2010) implementam um modelo para prever as taxas de câmbio utilizando a técnica de *Support Vector Regression* (SVR), comparando os resultados com as tradicionais Redes Neurais Artificiais (RNA). O modelo proposto pela técnica de (SVR) apresentaram melhor desempenho. Os erros médios quadrados de previsão são significativamente menores em comparado com as redes neurais artificiais.

Hong *et al.* (2011) enfatizam que realizar a previsão exata de fluxo de tráfego urbano tem sido uma das questões mais importantes a nível mundial na pesquisa sobre o congestionamento do tráfego rodoviário, a previsão de fluxo de tráfego envolve um padrão de dados bastante complexa não-linear. O modelo de *Support Vector Regression* (SVR) tem sido amplamente utilizado para resolver problemas de regressão não linear e de séries temporais. Para isso os autores apresentam um modelo de previsão de tráfego que utiliza o modelo de SVR para prever fluxo de tráfego. Além disso, um exemplo numérico de valores de fluxo de tráfego do norte Taiwan é empregado para elucidar a capacidade de previsão do modelo proposto. Os resultados de previsão indicam que o modelo proposto produz resultados de previsão mais precisos do que o sazonal auto-regressivo de média móvel integrado (SARIMA) modelo de séries temporais. Portanto, o modelo é uma alternativa promissora para a previsão de fluxo de tráfego.

Wu e Akabarov (2011) abordam a previsão do número de reclamações de garantia de produtos e serviços e ratificam que é de vital importância para os fabricantes / fornecedores de garantia a fim de preparar planos fiscais. Na literatura, um número de técnicas, tais como os modelos log-linear de Poisson, modelos de séries temporais e modelos de redes neurais artificiais foram desenvolvidos para preverem tais problemas. No entanto, pode-se encontrar dois pontos deficientes existentes nessas abordagens: (1) eles não consideram o fato de que garantia sinistros avisados nos últimos meses pode ser mais importante na previsão de garantia de futuro, e (2) são desenvolvidos com base em taxas de reparo, o que pode causar perdas de informações. Para superar estes dois pontos fracos, os autores apresentam duas abordagens diferentes para a previsão de reivindicações

de garantia, baseadas na metodologia do *Support Vector Regression* (SVR), utilizando tanto médias ponderadas como séries temporais. Com base na avaliação do modelo ao longo de seis meses antes de previsão, os resultados mostram que os modelos propostos apresentam um desempenho superior em comparação com a de perceptron multicamadas, baseado em Redes Neurais Artificiais (RNA).

Veronezi *et al.* (2011) utilizaram RNA no RP em radiografias de coluna lombar na incidência perfil para auxiliar no diagnóstico da osteoartrite primária com estudo transversal, descritivo, analítico, de abordagem quantitativa e com ênfase diagnóstica. O conjunto de treinamento foi composto por imagens coletadas de pacientes submetidos a radiografias digitais de coluna lombar na incidência em perfil.

No desenvolvimento de uma metodologia para a criação de etiqueta de qualidade com a finalidade de classificar elementos de uma mesma região e/ou grupo, utilizando o processo *KDD*, Góes (2012) fez uso das metaheurísticas Redes Neurais, *Support Vector Machine* e Algoritmos Genéticos, além de heurísticas baseadas em distâncias euclidianas, a fim de indicar a mais adequada para o problema em questão.

Wang *et al.* (2012) comentam a importância de prever a demanda de energia elétrica para uma indústria. Porém prever tal demanda torna-se difícil devido a influência de vários fatores que foram caracterizados como não-linear, o que demonstra que a previsão de carga anual não poderá ser solucionada pelos métodos lineares. Com isso os autores utilizaram a metodologia do *Support Vector Regression* (SVR) que demonstra ser útil no tratamento de problemas não-lineares de previsão. Porém os mesmos utilizaram um modelo híbrido que combina com algoritmo de evolução, onde o algoritmo é usado para escolher os parâmetros adequados para o modelo de previsão de carga (SVR). A eficácia deste modelo tem sido aprovado pela simulação final que mostra que o modelo proposto supera o modelo SVR com os parâmetros padrão, de volta a propagação de RNA (BPNN) e modelos de previsão de regressão por modelos lineares.

Mabilana *et al.* (2012) realizaram um trabalho como objetivo de ajustar um modelo matemático agrometeorológico espectral, para estimativa de produção do milho, na província de Manica em Moçambique, um país localizado ao longo da costa Leste da África. A área de estudo envolveu os distritos de Gondola, Manica, Mossurize e Sussundenga, responsáveis por mais de 80% da produção de milho na

província de Manica, nos anos de 2000 a 2009. O modelo agrometeorológico espectral, ajustado por meio de regressão linear múltipla, teve como variáveis independentes os índices meteorológicos e espectrais e, como variável dependente, o rendimento médio ou o rendimento relativo. O modelo regional, que incluiu os distritos de Gondola, Manica e Sussundenga, e considerou o rendimento relativo, foi o mais recomendado para estimativa de rendimentos do milho, na região, com $R^2 = 0,762$ e RMSE de 9,46%.

Sartori *et al.* (2012) ajustaram um modelo de regressão múltipla em função de atributos de imagens de radar polarimétrico, para estimativa da variável morfológica “volume da haste” de macrófitas encontradas na planície de inundação da Amazônia, no Lago Grande de Monte Alegre (Pará - Brasil). Com esta estimativa, pretende-se avaliar o potencial dos dados polarimétricos do sensor ALOS/PALSAR, destacando a importância da informação de fase das imagens de radar. Apesar do modelo não ter obtido um elevado coeficiente de determinação (44%), ele apresentou boa capacidade preditiva, já que todos os elementos de validação caíram dentro do intervalo de predição de 95% de confiança. Dentre as cinco variáveis independentes do modelo, quatro foram geradas a partir da informação de fase das imagens, o que reforça a importância dessa informação.

4 REFERENCIAL TEÓRICO

Este capítulo aborda os conceitos das técnicas aplicadas nesta dissertação, sua contextualização e as equações necessárias para o desenvolvimento da referida pesquisa.

4.1 DESCOBERTA DE CONHECIMENTO EM BASES DE DADOS (*KNOWLEDGE DISCOVERY IN DATABASES* - KDD)

De uma forma geral, todas as organizações possuem dados históricos armazenados em seus bancos de dados, com informações sobre clientes, fornecedores, compras e vendas, além de informações pertinentes a cada processo interno em particular. Podem ser empresas dos mais diversos ramos de produção, empresas de telecomunicações, instituições educacionais, hospitais, instituições financeiras, dentre tantas outras. Contudo apenas armazenar tais dados não é por si só um diferencial competitivo para a organização; é necessário, além disso, verificar e analisar tais dados armazenado em busca de informações relevantes, ou seja, verificar se há algum conhecimento a ser descoberto em tais dados.

A Descoberta de Conhecimento em Bases de Dados (*Knowledge Discovery in Databases*, ou simplesmente, *KDD*) é um processo que visa encontrar informações em bases de dados de uma maneira automatizada, criando relações de interesse que, muitas vezes, não são observadas por especialistas no assunto. (GÓES, 2012)

Uma das primeiras citações sobre a definição de *KDD* aconteceu no trabalho de Frawley, Piatetsky-Shapiro e Matheus (1991) como sendo “o processo, não trivial, de extração de informação, implícitas, previamente desconhecidas e úteis, a partir dos dados armazenados em uma base de dados”.

Outra definição do termo *KDD* surgiu três anos após no trabalho de Brachman e Anand (1994) que definiram como uma tarefa de descoberta de conhecimento intensivo, consistindo de interações complexas, feitas ao longo do tempo entre o homem e uma grande base de dados, possivelmente suportada por um conjunto heterogêneo de ferramentas.

Dois anos mais tarde Fayyad *et al.*, (1996) afirmam que o *KDD* é “o processo não trivial de identificação válida em dados novos, potencialmente úteis e finalmente com padrões compreensíveis”; é esta definição a mais comum utilizada na literatura.

Com esta última definição, pode-se definir alguns termos utilizados neste contexto, conforme apresentado no quadro 1 seguinte.

QUADRO 1 – Termos presentes na definição de *KDD*

TERMO	DESCRIÇÃO
Dados	Um conjunto de registros.
Padrão	Uma expressão E em uma linguagem L que descreve registros em um subconjunto dos dados.
Processo	Envolve várias etapas como a preparação de dados, busca de padrão, avaliação do conhecimento e refinamento com repetição após a modificação.
Válido	Padrões descobertos devem ser verdadeiros em novos dados com algum grau de certeza e generalizáveis no futuro para outros dados.
Novo	Os padrões devem ser novos, ou seja, não conhecidos previamente.
Útil	Os padrões levam a ações úteis.
Compreensível	O processo deve levar à percepção humana. Padrões devem ser transformados em conhecimento compreensível, a fim de facilitar uma melhor interpretação dos dados subjacentes.

FONTE: Góes (2012)

Dentre as etapas de um processo *KDD* encontra-se a de *Data Mining* (Mineração de Dados ou, simplesmente, *DM*). *DM* é a etapa principal, das cinco pertencentes ao *KDD*. As demais etapas são as seguintes, por ordem sequencial: seleção dos dados; limpeza dos dados ou pré-processamento; transformação dos dados; aplicação do *DM* e, finalmente, interpretação do conhecimento gerado. (FAYYAD *et al.*, 1996):

“[...] *KDD* refere-se ao processo global de descoberta de conhecimento útil a partir de mineração de dados que é uma etapa particular neste processo. A mineração de dados é a aplicação de algoritmos específicos para extrair padrões dos dados. [...] Os passos adicionais no processo de *KDD*, como preparação de dados, seleção de dados, limpeza de dados, a incorporação de conhecimento prévio adequado, e a interpretação adequada dos resultados da mineração, são essenciais para garantir que o conhecimento útil seja derivado dos dados. A aplicação direta de métodos da mineração de dados pode ser uma atividade perigosa, que leva facilmente à descoberta de padrões sem nexos e inválidos.” (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996)

A figura 12 apresenta a sequência da metodologia *KDD*, segundo Fayyad *et al.* (1996).

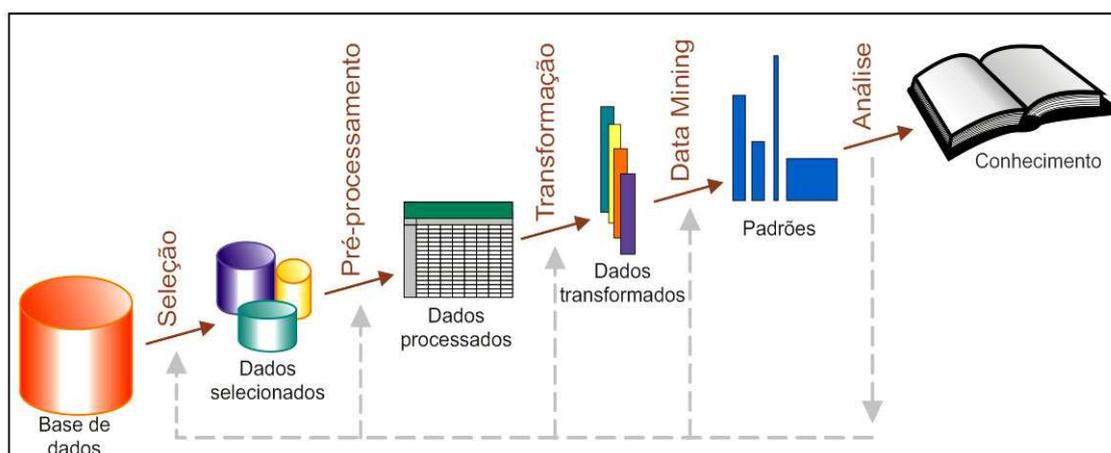


FIGURA 12 – Etapas do processo *KDD* por Fayyad *et al.*, 1996

FONTE: Góes (2012).

Feldens e Castilho (1997) e Bonchi *et al.* (2009) dividem a metodologia *KDD* apenas em três etapas: pré-processamento, mineração de dados e pós-processamento. Ao analisar as etapas três definidas, verifica-se que as três primeiras de Fayyad *et al.* (1996) estão inseridas na primeira de Feldens e Castilho (1997); o mesmo ocorre com última etapa definida por Fayyad *et al.* (1996) que está contemplada pela última de Feldens e Castilho (1997) e Bonchi *et al.* (2009).

Nesta dissertação, a sequência apresentada das etapas da metodologia de *KDD* terá os passos sugeridos por Fayyad *et al.* (1996). Durante este processo pode ser encontradas informações importantes já explícitas, ou não, pois tais informações podem ser já conhecidas/esperadas ou informações importantes que ao analisar a base de dados não se verificava nenhum tipo relação direta ou indireta.

Geralmente, tem-se como respostas informações que não se podem detectar quando se aplicam métodos tradicionais na análise de dados para posterior tomada de decisão, pois, em sua grande maioria, os métodos tradicionais são capazes de verificar apenas as relações explícitas nos bases de dados.

As etapas do processo *KDD*, constantes na Figura 12, são descritas a seguir.

A fase da **seleção dos dados** consiste na escolha do conjunto de dados que se pretende analisar, definindo assim os atributos e os registros. Em sua grande maioria, esta seleção é realizada por um especialista da área proveniente dos dados, pois possui papel fundamental no resultado final.

A **limpeza dos dados ou pré-processamento** é a fase em que se determinam quais dados serão eliminados por serem redundantes, ou por possuírem ruídos detectáveis e dados discrepantes dos demais. Além disto, é verificada a possibilidade de diminuir o número de variáveis. Para isto, podem ser aplicados métodos estatísticos, a fim de melhorar a eficácia dos algoritmos de classificação.

Na **transformação dos dados**, estes precisam ser armazenados e formatados de forma adequada à aplicação do algoritmo na próxima fase. Também é nesta fase que são determinados atributos faltantes que podem ser obtidos de outros como, por exemplo, a duração de certo evento por meio do horário inicial e horário final da ocorrência do mesmo.

Na etapa de **Data Mining**, a mais importante do processo *KDD*, são aplicadas as técnicas para análise dos dados por meio de heurísticas ou meta-heurísticas, para a descoberta de padrões. O tempo de execução desta fase deve ser compatível com o tempo disponível na espera da solução. Muitos são os métodos utilizados nesta etapa, tais como Redes Neurais Artificiais, *Support Vector Machine*, Regressão Múltipla, Árvores de Decisão, Técnicas Estatísticas, dentre outras.

Na **interpretação do conhecimento** gerado deve-se interpretar o conhecimento apresentado, verificando a relevância ou não dos padrões e, com isso, verificar também a eficácia do método aplicado na etapa do *DM*. Caso o analista julgue que o conhecimento não é válido, o processo deverá ser reiniciado, analisando-se todas as etapas em busca de melhorar e/ou refazer o que for necessário até que o conhecimento obtido seja julgado como verdadeiro por quem o analisa.

4.1.1 DATA MINING (DM OU MINERAÇÃO DE DADOS)

O *DM* é a etapa mais importante do processo *KDD*, conforme já comentado, pois é nesta etapa que se aplicam técnicas para análise dos dados, seja através de procedimentos heurísticos ou meta-heurísticos, para a descoberta de padrões (WITTEN e FRANK, 2005). Pode-se dizer também que a Mineração de Dados é um processo de descoberta automática de novos e compreensíveis modelos e padrões em grandes quantidades de dados (ZAKI, KARYPIS E YANG 2007).

Neste contexto Côrtes, Procaro e Lifschitz (2002) afirmam que este é um processo altamente cooperativo entre homens e máquinas, visando à exploração de

grandes bases de dados, com o objetivo de extrair conhecimentos através do RP relacionamento entre as variáveis, conhecimentos esses que possam ser obtidos por técnicas conhecidas, como apresentados nesta dissertação, e validados por sua expressividade estatística.

Ao analisar e explorar o banco de dados, Thuraisingham *et al.* (2005) afirmam que se realizam várias consultas e extração de informações úteis e que, muitas vezes, são desconhecidas e inesperadas, como padrões e tendências. No entanto, a busca em grandes bases de dados só é possível devido aos métodos computacionais e a análise do ser humano (WEISS E INDURKHYA, 1998).

De forma mais simples, Han e Kamber (2006) definem *DM* como o processo de extração ou mineração de conhecimento em grandes quantidades de dados. Esta fase não elimina o conhecimento específico da área de origem dos dados, mas sim, auxilia os analistas a encontrar os padrões nas bases de dados. O *DM* tem como base a Estatística, a Inteligência Artificial, a Aprendizagem de Máquinas, o RP e os Sistemas de Bases de Dados. É adequada à análise de grandes quantidades de dados, dados com alta dimensionalidade e com natureza distribuída e heterogênea (TAN, STEINBACH E KUMAR, 2005).

Além disso, estes analistas devem verificar e assegurar se os resultados são significativos para a solução do que se pretende. São eles que podem determinar os dados relevantes para o processo.

A aplicação deste processo tem como principais resultados, a obtenção de: associação; agrupamento; previsão; classificação de padrões; dentre outros. A *classificação de padrões* é a tarefa que generaliza uma estrutura conhecida a ser aplicada aos novos dados, ou seja, procura descobrir uma função a ser aplicada aos novos dados que consiga classificar registros em um conjunto de dados pré-definidos.

O agrupamento (ou *clusterização*) de padrões consiste em agrupar registros em subconjuntos que de alguma forma compartilhem propriedades, onde os grupos variáveis não são pré-definidos.

A *associação de padrões* busca relações entre as variáveis do problema, sendo que estas podem ser independentes ou exploratórias. Por ser baseado em Inteligência Artificial, com apoio da Matemática e/ou Estatística, têm-se diversas técnicas apresentadas na literatura para a realização de *DM*, como as já mencionadas anteriormente.

4.2 PROCESSO DE CLASSIFICAÇÃO

Para Castro e Prado (2002), a classificação de padrões pode ser definido como um procedimento em que se busca a identificação de certas estruturas nos dados de entrada em comparação a estruturas conhecidas e sua posterior dentro de categorias, de modo que o grau de associação seja maior entre estruturas de mesma categoria e menor entre as categorias de estruturas diferentes.

Procedimentos de classificação de padrões podem ser utilizados no processamento de documentos como leitura de crachás com código de barras e reconhecimento dos caracteres e algarismos em placas de veículos, sinal da voz, assinaturas, códigos de barras, impressão digital, e outros (THEODORIDIS e KOUTROUMBAS, 2008).

Para Ales (2008), os padrões podem simplesmente ser encontrados em objetos como, por exemplo, um lápis. Algumas de suas características são: um grafite ou giz colorido, em formato de prima ou cilindro revestido por um invólucro de madeira, uma extremidade apontada utilizada para escrever ou desenhar, comprimento de aproximadamente quinze centímetros, cores diversificadas e inúmeras outras características que constituem seus padrões.

Castro e Prado (2002) e Ales (2008) enfatizam que na execução de um sistema o RP necessita englobar três grandes etapas, sendo elas:

- a) Representação dos dados de entrada e sua mensuração;
- b) Extração das características;
- c) Identificação e classificação do objeto em estudo.

A primeira etapa refere-se à representação dos dados de entrada que podem ser mensurados a partir do objeto a ser estudado. O vetor que caracteriza perfeitamente um objeto seria de dimensionalidade infinita, descrito por um vetor X , conforme a seguir.

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$$

onde: $x_1, x_2, x_3 \dots x_n$ são suas características (ou atributos).

A segunda etapa consiste na seleção dos atributos, ou seja, na extração de características intrínsecas e atributos do objeto e consequente redução da dimensionalidade do vetor padrão. A escolha das características é de fundamental importância para um bom desempenho da técnica de RP para realizar a classificação de padrões e deve ser realizada objetivando os fenômenos que se pretende classificar, ou seja, é exigido um conhecimento específico sobre o problema em estudo. Nesta etapa, os objetivos básicos são: a redução da dimensionalidade do vetor característico, sem que isso implique em perda de informação que possa ser relevante para a classificação, objetivando a redução do esforço computacional e a seleção das características significativas para a tarefa de RP.

A terceira etapa no processo de classificação de padrões envolve a determinação de procedimentos que possibilitem a identificação e classificação do objeto em uma classe de objetos. De modo diferente da segunda etapa, aqui a concepção do classificador pode ser abordada de forma abstrata e independente da natureza do problema.

Em um sistema físico qualquer existe um número infinito de características que definem os padrões nele existentes. A metodologia de RP tem como função identificar, determinar e extrair as características mais significativas do banco de dados que representem a descrição do objeto estudado, dentre as infinitas características que possam descrevê-lo. Outro dado relevante é que o extrator de características varia com o sistema a ser analisado, ou seja, dependerá do objeto em estudo, sendo que para cada caso estudado existem suas características particulares, seus padrões.

A tabela 2 exemplifica várias tarefas de classificação, propostas por um sistema de RP, com seus dados de entrada e respectivos dados de saída:

TABELA 2: Exemplos de classificação, uma das tarefas do RP.

Tarefas de Classificação	Dados de Entrada	Dados de Saída
Desmatamento Florestal	Imagens de satélite	Identificação da área desmatada
Reconhecimento de Voz	Dados da Voz	Identificação da palavra
Diagnósticos Médicos	Sintomas	Identificação da Patologia
Previsão do tempo	Mapas atmosférico:	Chuva, Sol etc.
Concessão de crédito	Dados dos clientes	Conceder ou não o crédito

FONTE: Autor (2014)

Uma vez extraídas as características é necessária a classificação do objeto, que pressupõe a designação do objeto a uma determinada classe. Nesta etapa o classificador “aprende” a distinguir dentre as classes, aquela a qual o objeto pertence, aglomerando padrões de uma mesma classe em agrupamentos. Se o treinamento do classificador exigir amplo conhecimento “a priori” da estrutura estatística dos padrões a serem analisados e o padrão de entrada for identificado como membro de uma classe pré-definida pelos padrões de treinamento, o classificador será chamado de Classificador Paramétrico e a classificação se processa de forma supervisionada (CASTRO E PRADO, 2002).

Uma técnica de classificação de padrões deve basear-se no conhecimento extraído dos exemplos de uma base de dados, com o objetivo de classificar um exemplo novo, desconhecido até então, à classe que mais reflita as suas características (CARVALHO e BRAGA, 2004).

A grande dificuldade na implementação de um sistema de classificação de padrões está justamente na escolha das técnicas mais adequadas para que as três fases anteriormente apresentadas ocorram de modo a representar satisfatoriamente os fenômenos do mundo real. A figura 13 ilustra, de forma mais detalhada, as diversas fases de um problema de classificação de padrões.

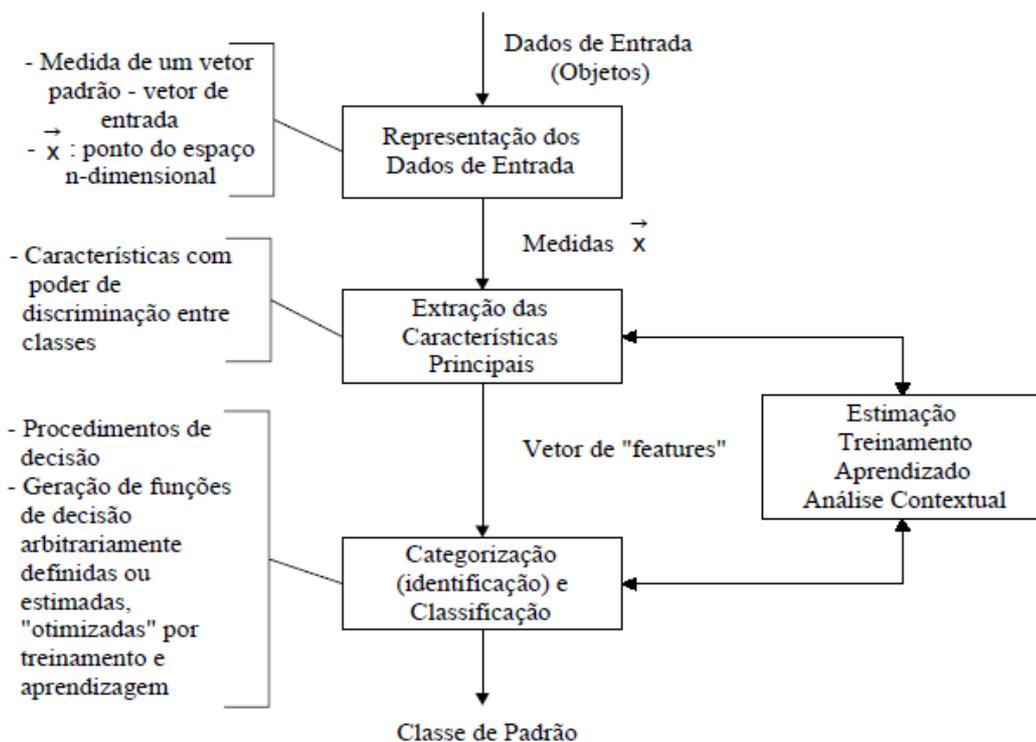


FIGURA 13– Fases para a Classificação de Padrões, uma das tarefas do RP.

FONTE: Castro e Prado (2002).

4.3 PROCESSO DE PREVISÃO

Um processo formal de previsão inicia-se com a definição clara do objetivo principal que está associada à necessidade da previsão. Por exemplo, a finalidade pode ser a previsão de uma demanda para reposição de estoques; a previsão para o planejamento da capacidade de uma nova linha de produto ou simplesmente o que poderá ocorrer em um processo produtivo (LUSTOSA *et al.*, 2008). Os mesmos autores enfatizam que dependendo do objetivo determinam-se, dentre outros aspectos, o nível de precisão da previsão, quais variáveis devem participar do processo, quais os períodos a serem previstos, quais os recursos envolvidos, entre outros.

Krajewski, Ritsman e Malhotra (2010) enfatizam que em um processo de previsão, o tomador de decisão deverá, antes de mais nada, planejar suas ações, a fim de minimizar os erros de previsão. Ele deve tomar três decisões iniciais: (1) o que deve ser previsto; (2) que tipo de técnica utilizar; e (3) qual modelo de previsão utilizar.

A partir da definição dos objetivos e da abrangência do processo de previsão, Lustosa *et al.* (2008) afirmam que o passo seguinte consiste na coleta dos dados, que inclui os registros históricos dos dados e as informações sobre os eventos do processo pesquisado, que podem indicar a necessidade de correção ou mesmo exclusão de alguns dados da série histórica. Evidentemente, a quantidade e o tipo das informações estão associados ao propósito da previsão, assim como a natureza do modelo de previsão a ser empregado.

O processo de previsão é uma metodologia muito importante nas empresas, serve para programar a utilização das máquinas de maneira adequada, para realizar a reposição dos materiais no momento e na quantidade certa, e para que todas as demais necessidades do processo industrial sejam adequadamente programadas. Apesar de as previsões serem importantes e úteis para o julgamento das atividades, elas apresentam erros em suas estimativas, devendo-se tomar cuidado tanto na coleta dos dados, quanto na escolha do modelo de previsão, com objetivo de minimizar seus erros (MARTINS E LAUGENI, 2005).

Para implementar um processo de previsão, além da base de dados e dos recursos computacionais, a organização necessita dos modelos de previsão. Lustosa *et al.* (2008) afirmam que de forma geral, esses modelos podem ser classificados em qualitativos e quantitativos. No primeiro grupo estão as técnicas

baseadas em consenso e opiniões, principalmente dos especialistas do processo estudado. Por outro lado, métodos quantitativos caracterizam-se pelo emprego de técnicas baseadas em pesquisa operacional que podem ser técnicas estatísticas, processos estocásticos, meta-heurísticas, dentre outras.

Krajewski, Ritsman e Malhotra, (2010) corroboram com essa informação enfatizando a existência de dois tipos gerais das técnicas de previsão para prever os resultados de um processo ou a demanda de um produto, os métodos qualitativos e os quantitativos. Os métodos qualitativos convertem as opiniões dos gerentes e especialistas em estimativas quantitativas, enquanto que os métodos quantitativos usam dados históricos das variáveis independentes do comportamento do processo a ser estudado para então projetar e prever o que possivelmente poderá acontecer nos eventos futuros.

A previsão qualitativa é claramente necessária quando não há dados quantitativos disponíveis para usar uma abordagem de previsão quantitativa. Contudo as abordagens de avaliação qualitativa podem ser usadas com abordagens quantitativas para melhorar a qualidade das previsões.

Os métodos quantitativos são possíveis apenas quando existem dados históricos adequados, muitas vezes chamados de arquivos históricos. Entretanto tais arquivos históricos podem ser inexistentes quando o processo está em sua fase inicial ou quando a empresa não julga necessário armazenar estas informações. Os arquivos históricos podem existir, porém podem ser menos úteis quando os valores arquivados não sejam precisos ou confiáveis, isso pode ocorrer quando os equipamentos apresentam falhas na leitura dos dados por uma imprecisão de leitura, ou simplesmente um equipamento defeituoso (KRAJEWSKI, RITSMAN E MALHOTRA, 2010).

Atualmente, as empresas dispõem de tecnologias de informação para organização das informações dos dados históricos do processo a ser analisado, facilitando o processo de coleta de dados para auxiliar os modelos quantitativos.

Notas de esclarecimento nos arquivos históricos são de fundamental importância para compreender certos padrões de comportamento e possíveis ajustes nos dados.

Krajewski, Ritsman e Malhotra, (2010) afirmam que tais anotações podem explicar eventos como problemas em equipamentos, elevação momentânea da produção ou até mesmo uma parada abrupta no processo. Estes eventos, dentre

outros, podem ocasionar o arquivamento de informações que fogem do padrão habitual do processo em si, podendo ocasionar erros na interpretação dos dados e previsões equivocadas.

Uma vez definido o modelo, Lustosa *et al.* (2008) complementam afirmando que o passo seguinte consiste na realização dos testes para aferir e avaliar a confiabilidade das previsões geradas, ou seja a validação do modelo. A aferição consiste, a partir da simulação com dados históricos, na determinação dos valores dos parâmetros do modelo em uso que produzem o menor erro de previsão.

Finalmente, uma vez validado o modelo e disponibilizados os recursos e as informações necessárias para a previsão, este pode ser incorporado à rotina operacional. Essa etapa consiste na implementação efetiva do processo de previsão. Uma vez implementado, o modelo deve ser monitorado e revisto periodicamente, com objetivo de preservar a qualidade das informações produzidas (LUSTOSA *et al.* 2008). O procedimento completo de um processo de previsão pode ser visualizado na figura 14.

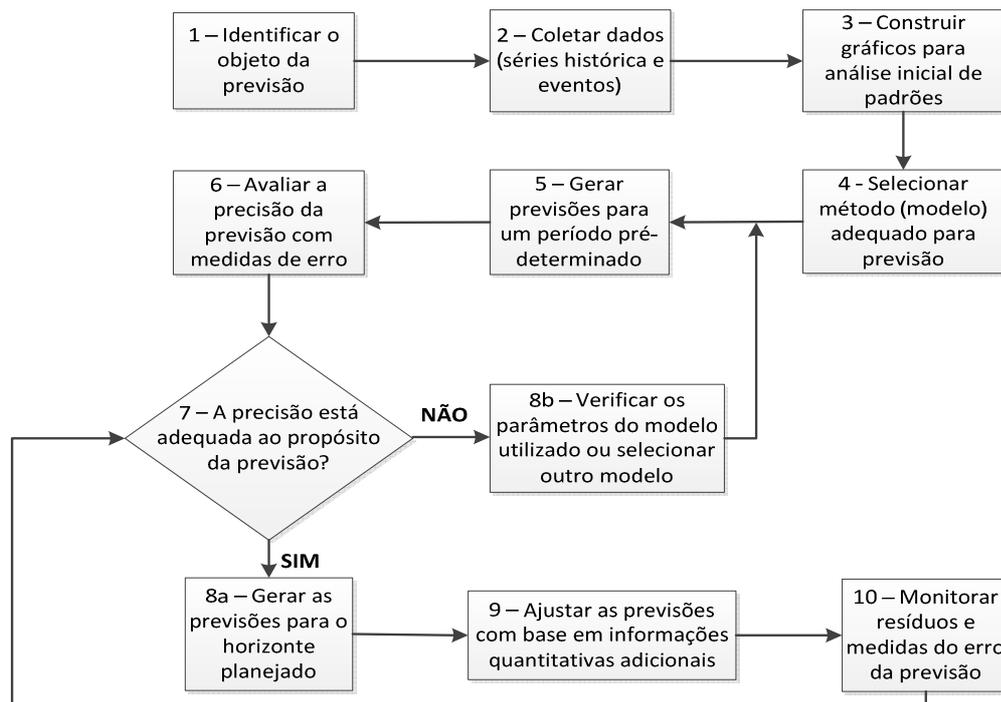


FIGURA 14 – Fluxograma de um processo de previsão

FONTE: Lustosa *et al.* (2008)

Martins e Laugeni (2005) ressaltam que antes de iniciar um processo de previsão deve-se, primeiramente, compreender a diferença entre predição e

previsão. Os autores enfatizam que a predição é um processo lógico que determina um acontecimento futuro, baseado em dados subjetivos e sem uma metodologia clara, enquanto que a previsão está baseada em processos metodológicos para a determinação de dados futuros baseados em um modelo estatístico, matemático ou econométrico ou ainda modelos subjetivos apoiados em uma metodologia de trabalho clara e previamente definida.

Vale ressaltar que o processo de previsão não é uma atividade independente ou isolada dos demais processos, mas, ao contrário, ela faz parte de um processo maior que inclui, além da previsão, o processo decisório, ou seja, o processo de tomada de decisão. De nada adiantará realizar uma previsão assertiva se não souber o que fazer com esse resultado.

4.4 TÉCNICAS APLICADAS NO ESTUDO

Nesta seção abordam-se as técnicas utilizadas nesta pesquisa para o processo de RP (RNAs e SVM).

4.4.1 REDES NEURAIS ARTIFICIAIS

O primeiro modelo de uma RNA baseada em um neurônio biológico teve início em 1943 através do trabalho pioneiro de Warren McCulloch e Walter Pitts, um psiquiatra e neuroanatomista e o outro um matemático recém-graduado. Juntos publicaram um trabalho com ideias sobre redes lógicas de nodos (neurônios) e os métodos de aprendizagem para que estes pudessem ser capazes de executar uma função, se prendendo mais na descrição de um modelo artificial de um neurônio biológico e suas capacidades computacionais (BRAGA *et al.*, 2011).

A partir deste estudo outros pesquisadores demonstraram seus trabalhos nesta área, como Hebb que em 1949, mostrou como a aprendizagem de redes neurais artificiais pode ser conseguida pela variação dos pesos de entrada dos neurônios.

O modelo PERCEPTRON surgiu em 1958 por Frank Rosenblatt, que tinha como topologia de RNA, uma estrutura de ligação entre os neurônios em forma de três camadas. A primeira com o propósito de receber os dados. A segunda recebendo os dados da primeira através das conexões com seus respectivos pesos

ajustáveis, enviando os sinais para a terceira camada de saída em forma de resposta. Pelo ajuste dos pesos o PERCEPTRON pode ser treinado de acordo com as informações recebidas do banco de dados durante a etapa de treinamento.

As redes neurais do tipo PERCEPTRON eram limitadas para a solução de alguns problemas que não demonstravam linearidade que eram difíceis de aprender, representando uma grande classe de funções com os algoritmos de aprendizagem desenvolvidos até então, o que culminou em uma freada nos trabalhos de exploração das RNA. Com o desenvolvimento do algoritmo de retropropagação (*backpropagation*) em 1982 por John Hopfieldo (BRAGA *et al.*, 2011), reacendeu-se o interesse pelos estudos das redes neurais artificiais pelo fato deste algoritmo proporcionar treinamento para as redes serem capazes de resolver problemas não lineares.

As RNAs podem ser definidas como ferramentas computacionais que utilizam modelagem matemática baseada nos conceitos do funcionamento dos neurônios dos seres humanos. Segundo Haykin (2005), uma RNA pode ser vista como uma máquina adaptativa com capacidade de aprendizagem, graças à interligação maciça de células computacionais denominadas neurônios. O mesmo autor afirma que uma RNA pode ser definida ainda como “um processador maciçamente paralelamente distribuído constituído de unidades de processamento simples, que tem a propensão natural para armazenar conhecimento experimental e torná-lo disponível para uso”.

Uma RNA funciona baseada na estrutura do cérebro humano que é responsável pela execução das funções sensoras, motoras e autônomas do corpo, além das emoções, pensamento e percepção. O cérebro com suas redes de neurônios tem ainda a capacidade de reconhecer padrões, armazenar conhecimento através de experiências e interpretar observações. Os neurônios biológicos são divididos em três partes, o corpo da célula, os dendritos e o axônio, cada um com uma função específica. Os dendritos tem a função de receber as informações através de impulsos nervosos vindos de outros neurônios e leva-los até o corpo celular, onde a informação é processada gerando novo impulso. Este impulso é transmitido através do axônio até os dendritos do neurônio seguinte (BRAGA *et al.*, 2011).

Esta ligação entre o axônio de um neurônio e o dendrito do outro é chamada de conexão sinapse. Esta união funcional entre os neurônios pela sinapse é que

formam as redes neuronais do nosso cérebro. As sinapses funcionam como um controlador de fluxos de impulsos entre os neurônios da rede de forma variável, sendo esta variação que confere ao neurônio a capacidade de adaptação. A figura 15 representa a configuração da comunicação entre neurônios biológicos (BRAGA *et al.*, 2011).



FIGURA 15 – Representação da comunicação entre neurônios biológicos

FONTE: BRAGA *et al.* (2011)

As unidades de processamento denominadas neurônios artificiais são programadas para realizar cálculos matemáticos baseados em funções. Segundo Braga *et al.* (2011), estas unidades são dispostas em uma ou mais camadas interligadas por um grande número de conexões normalmente unidirecionais. Um neurônio artificial é uma versão simplificada de um neurônio biológico, que utiliza modelagem matemática para realização das funções. Eles possuem n terminais de entrada e um terminal de saída. Para simular o efeito da sinapse são atribuídos pesos na conexão de entrada que podem ter valores variados. Para a ativação do neurônio é utilizada uma função que ativa ou não na sua saída, dependendo dos valores obtidos com o resultado desta função. A figura 16 ilustra a descrição de um neurônio artificial.

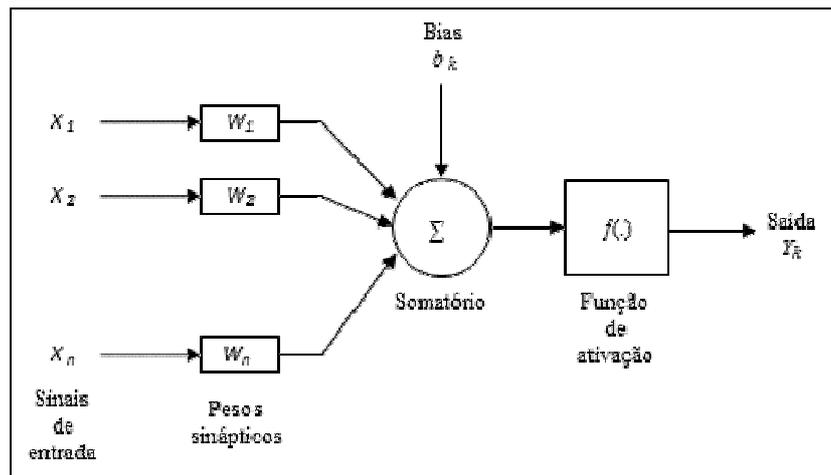


FIGURA 16 – Modelo de um neurônio artificial

FONTE: Haykin (2005).

Sendo os **Sinais de entrada** representados por X_1 , X_2 e X_n conectados ao neurônio são multiplicados pelos pesos sinápticos W_1 , W_2 e W_n respectivo a cada sinal de entrada. Os **Pesos sinápticos** podem assumir valores negativos e são responsáveis pelo ajuste dos dados de entrada para correção dos erros no valor de saída. O **Somatório** realiza o somatório dos sinais de entrada, ponderados pelas respectivas sinapses. Ele conta ainda com a entrada de um bias que tem o efeito de aumentar ou diminuir a entrada da função de ativação. A **Função de ativação** tem por finalidade restringir a amplitude da saída de um neurônio em um intervalo de valor finito, normalmente com valor unitário $[0,1]$ ou $[-1,1]$ determinando a relação entre a entrada e saída do neurônio.

4.4.1.1 TIPOS DE REDES NEURAIIS

A estruturação de uma RNA está relacionada com o tipo de algoritmo de treinamento que vai ser usado. De acordo com a estrutura de ligações dos neurônios de uma RNA, têm-se tipos diferentes de arquiteturas. Segundo Haykin (2005), existem três classes de arquiteturas de redes diferentes. São elas: Redes de camada única, redes de múltiplas camadas e redes recorrentes.

De acordo com Braga *et al.* (2011), a definição da estrutura de uma RNA é muito importante, pois a estrutura é um fator que restringe o tipo de problema que pode ser tratado pela rede, que é definido pelo número de camadas da rede, número de neurônios em cada camada, tipo de conexão entre os neurônios e a topologia da rede. Temos então as redes de camada única e as redes de múltiplas camadas.

As redes de camada única são formadas por uma camada de entrada de nós e uma camada de saída de neurônios. A camada de entrada projeta sobre a camada de neurônios em sentido único. A camada de entrada não é contada para a definição do número de camadas, porque nesta, não é realizada nenhuma ação computacional. Braga *et al.* (2011) enfatizam que as redes de camada única só são capazes de resolver problemas linearmente separáveis. Nas redes de camada única só existe um nó entre qualquer entrada e qualquer saída da rede.

Entretanto Campos e Saito (2004) enfatizam que também existem as RNA's multicamadas com o algoritmo de retroprogramação (*backpropagation*) para a aprendizagem. Seu desenvolvimento alavancou o uso das redes neurais artificiais por permitir a sua utilização em estruturas mais complexas, que são capazes de resolver problemas que não podiam ser tratados por outros tipos de redes neurais. Este tipo de RNA tem boa capacidade para áreas como: Controle, RP, estatística e outras.

Estas redes neurais são normalmente chamadas de PERCEPTONS de múltiplas camadas (MLP), e tem sido aplicada com sucesso para resolver problemas difíceis através de treinamento de forma supervisionada com o algoritmo *backpropagation*.

As redes de multicamadas são formadas por mais de uma camada de neurônios entre alguma entrada e uma saída da rede. Ela “se distingue pela presença de uma ou mais camadas ocultas, cujos nós computacionais são chamados correspondentemente de neurônios ocultos ou unidades ocultas” (HAYKIN, 2005).

A camada oculta tem a função de interagir entre a camada de entrada e a camada de saída da rede. Quanto maior o número de camadas ocultas, maior a capacidade de processamento da rede. Estas camadas ocultas não fazem parte da entrada ou da saída da RNA, por isso são denominadas ocultas.

A figura 17 mostra a estrutura de uma rede multicamada com uma camada oculta. Pode-se observar que a rede é totalmente conectada, ou seja, um neurônio em qualquer camada da rede esta conectado a todos os neurônios da camada anterior.

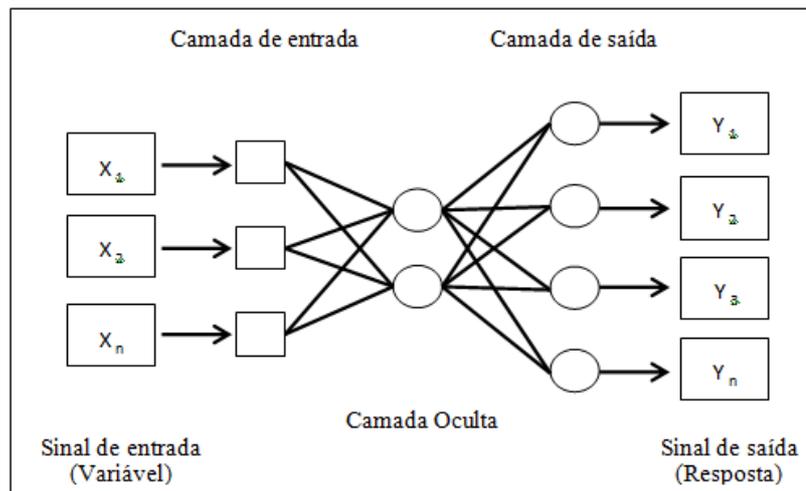


FIGURA 17– RNA com uma camada oculta.

FONTE: Adaptado de Haykin (2005).

Segundo Haykin (2005), os sinais de entrada se propagam para frente através da rede, camada por camada. Os nós da fonte da camada de entrada fornecem os padrões de ativação para os neurônios da segunda camada. Os sinais de saída da segunda camada são utilizados como entrada para a terceira camada, e assim sucessivamente para o restante da rede. O conjunto de sinais de saída da camada final de neurônios da rede forma a resposta para os padrões de entrada na primeira camada de nós de ativação.

4.4.1.2 TIPOS DE APRENDIZAGEM DE UMA REDE NEURAL

Uma das características que chama a atenção nas Redes Neurais é a capacidade de aprendizagem, que é realizada através de treinamento, tomando como base os valores das variáveis de entrada e os resultados das variáveis respostas. “A capacidade de aprender através de exemplos e de generalizar a informação aprendida é, sem dúvida, o atrativo principal da solução de problemas através das RNA's.” (BRAGA *et al.*, 2011).

O aprendizado de uma RNA através do ambiente que ela está inserida pode ser o seu principal atributo. Esta habilidade de aprender dentro do seu próprio ambiente pode melhorar o seu desempenho.

“A aprendizagem é um processo pelo qual os parâmetros de uma RNA são adaptados através de um processo de estimulação pelo ambiente no qual a rede está inserida. O tipo de aprendizagem é determinado pela maneira pela qual a modificação dos parâmetros ocorre.” (HAYKIN, 2005).

O processo de aprendizagem de uma RNA inicia-se pela formação do banco de dados, que de acordo com Campos e Saito (2004), deve ser dividido em dois conjuntos, sendo: um conjunto para o treinamento, que deve corresponder de 80 a 90% dos dados, e um segundo conjunto para testes e validação da rede com 10 a 20% dos dados. O banco de dados para o treinamento da rede deve ter tamanho suficiente para conter todas as condições possíveis do processo operacional para garantir que o treinamento da rede seja abrangente. O mesmo critério deve ser observado com o conjunto de dados para a avaliação da rede.

A figura 18 representa o fluxograma de processo de aprendizagem utilizado para a configuração da RNA em estudo.

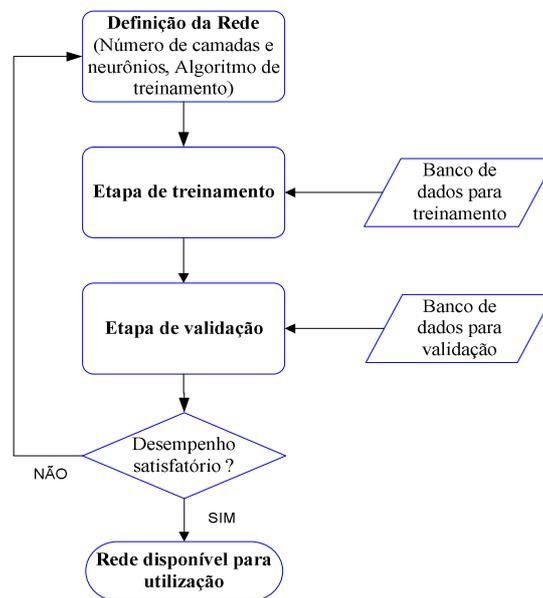


FIGURA 18 - Fluxograma do processo de aprendizagem da RNA

FONTE: Adaptado de Braga *et al.* (2011).

Foram desenvolvidos alguns métodos de treinamento de uma RNA, que usam como ferramentas, diversos tipos de algoritmos de aprendizagem. De acordo com a modelagem matemática do algoritmo a aprendizagem pode ser dividida em dois grupos distintos, Aprendizagem Supervisionada e Aprendizagem não Supervisionada.

Na aprendizagem não supervisionada não conta com a existência de um professor para acompanhar o treinamento da rede. Neste tipo de treinamento a rede só conta com as variáveis de entrada. Segundo Braga *et al.* (2011), a partir do momento que a rede estabelece harmonia com os dados de entrada de forma estatística, ela consegue desenvolver a habilidade de formar internamente,

representações que codificam as características das variáveis de entrada, criando assim automaticamente novos grupos. Para que seja possível este tipo de treinamento é necessário que os dados de entrada sejam redundantes de tal forma que a rede possa sempre comparar os dados entre si.

Já na aprendizagem supervisionada, que tem esta definição em função das variáveis de entrada e de saída serem fornecidas por um supervisor, que tem a função de monitorar a resposta da RNA para cada valor de entrada. O treinamento é feito com base nos pares de dados de entrada e saída. Este tipo de entrada permite à rede um ajuste dos pesos, uma vez que os valores da resposta para cada dado de entrada é conhecido, permitindo assim, uma aproximação com o valor desejado dentro de um limite de erro estabelecido (REZENDE, 2005).

Braga *et al.* (2011), afirmam que este é o tipo de treinamento mais comum, tanto de neurônios com pesos quanto os sem pesos. O objetivo é ajustar os parâmetros de tal forma que se possa encontrar uma ligação entre os pares de entrada e de saída. A desvantagem do aprendizado supervisionado é que na ausência do professor a rede não conseguira aprender novas soluções para situações que não estão nos exemplos de treinamento. A figura 19 ilustra o processo de aprendizagem supervisionada.

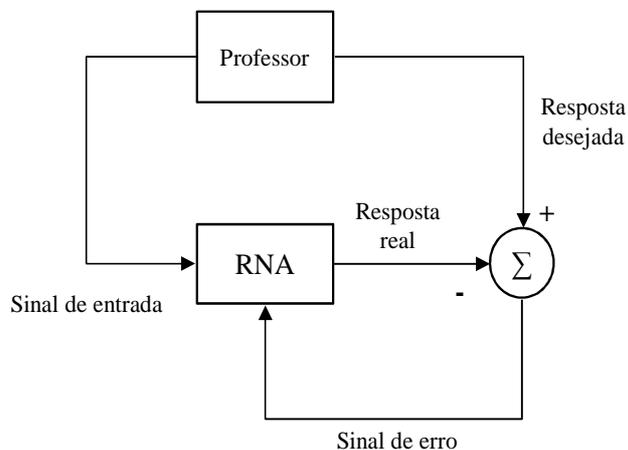


FIGURA 19 – Representação da aprendizagem supervisionada

FONTE: Adaptado de Braga *et al.* (2011).

4.4.1.3 ALGORITMO *BACKPROPAGATION*

O algoritmo de treinamento para redes neurais *backpropagation* surgiu em meados da década de 1980, criado por Rumelhart, Hinton e Williams. Segundo

Rezende (2005), o mesmo veio solucionar o problema de treinamento das redes de multicamadas, que são usadas para resolução de problemas com maior complexidade.

“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” (BRAGA *et al.*, 2011).

O treinamento da rede é feito através da avaliação dos erros medidos com a comparação entre os valores de saída desejados e o calculado pela rede. De posse dos resultados dos erros o algoritmo realimenta a camada oculta para que os pesos atribuídos às variáveis sejam alterados. Este processo se repete até que os pesos sejam ajustados de tal forma que o erro seja o menor possível.

De acordo com Braga *et al.* (2011), o treinamento acontece em duas fases. A primeira fase, *forward*, que é utilizada para definir a saída da rede utilizando o padrão de entrada, e a segunda fase, *backward*, que utiliza a saída informada e a saída fornecida pela rede para atualização dos pesos. O fluxo desse processamento ocorre em sentidos contrários. Na *forward* os dados das variáveis de entrada seguem para a camada de saída. Na *backward* os erros seguem da saída para a entrada das conexões dos neurônios.

O que caracteriza o treinamento com o algoritmo *backpropagation* segundo Rezende (2005), é o ajuste dos pesos das camadas intermediárias através da retroalimentação, ou seja, propagação para traz do erro medido na camada de saída.

4.4.2 SUPPORT VECTOR MACHINE (SVM)

Desenvolvido por Boser, Guyon e Vapnik, o algoritmo do *Support Vector Machine* (SVM) foi apresentando pela primeira vez em 1992 na *Annual Workshop on Computational Learning Theory*. O seu nome enfatiza a importância que os vetores mais próximos da margem de separação representam, uma vez que eles determinam a complexidade do vetor de suporte (CARVALHO, 2005). Em sua forma inicial, também denominada *Support Vector Classification* (SVC), o algoritmo propõe encontrar por meio da metodologia de aprendizagem um hiperplano separador de

margem máxima (SUNG e MUKKAMALA, 2003; DING e DUBCHAK, 2001; ALES, 2008).

Em 1995, Corinna Cortes e Vapnik sugeriram uma modificação no algoritmo, introduzindo variáveis de folgas para obter uma margem suave no hiperplano separador (CORTES e VAPNIK, 1995), minimizando a probabilidade de classificação errada dos padrões.

Tal hiperplano permite a classificação de dados linearmente separáveis, seja no espaço de entrada ou em um espaço característico de dimensão maior. A classificação neste último espaço é permitida pelo uso de funções de transformação *Kernel* (BELTRAMI, 2009).

Góes (2012) enfatiza que o SVM foi desenvolvido para a classificação binária dos elementos, ou seja, dado o conjunto de treinamento, onde as entradas são representadas por $x_i \in R^n$ e as respectivas saídas por $y_i \in \{0,1\}$ ou $\{-1,1\}$.

Na figura 20 ilustra-se o referido conjunto de dados, na qual se identificam duas classes distintas denominadas classe -1 e classe +1, ou classe negativa e positiva respectivamente.

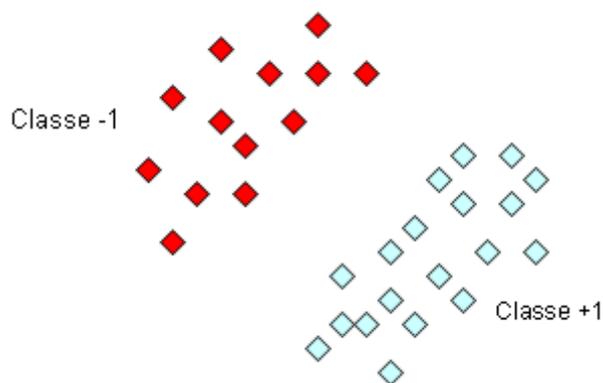


FIGURA 20 - Conjunto de dados linearmente separáveis

FONTE: Ales (2008)

Na figura 21 observa-se que existem várias formas de separar a amostra de dados, sem erros.

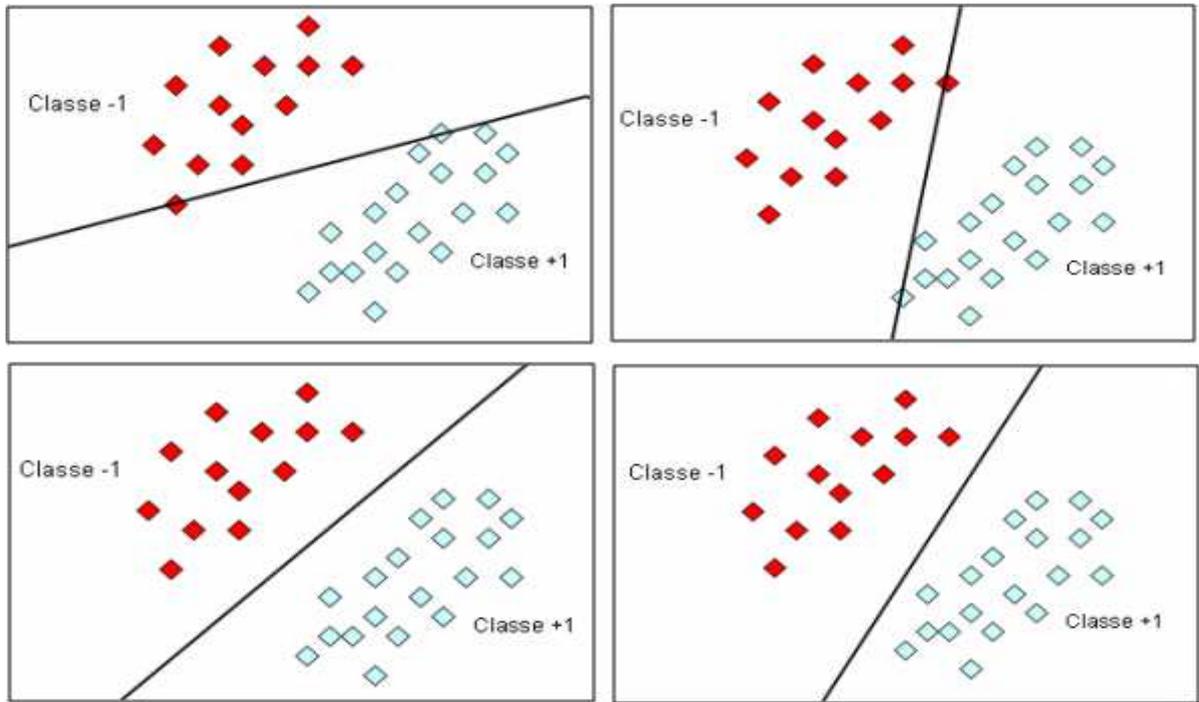


FIGURA 21 - Possibilidades de separação por meio de diferentes hiperplanos

FONTE: Beltrami (2009)

No entanto, há somente uma maneira na qual se maximiza a margem de classificação que é definida como a distância entre hiperplano separador e o vetor mais próximo de cada classe (BELTRAMI, 2009; LIMA, 2004). O hiperplano que maximiza a distância entre estas duas margens é conhecido como hiperplano ótimo.

Resultados teóricos apresentados em Vapnik (1998, *apud* Carvalho, 2005) indicam que a maximização das margens entre os exemplos das classes -1 e +1, implica uma maior generalização do sistema de aprendizagem. Com isso, pode-se dizer que o objetivo do SVM é encontrar esse hiperplano particular, o qual está ilustrado na figura 22.

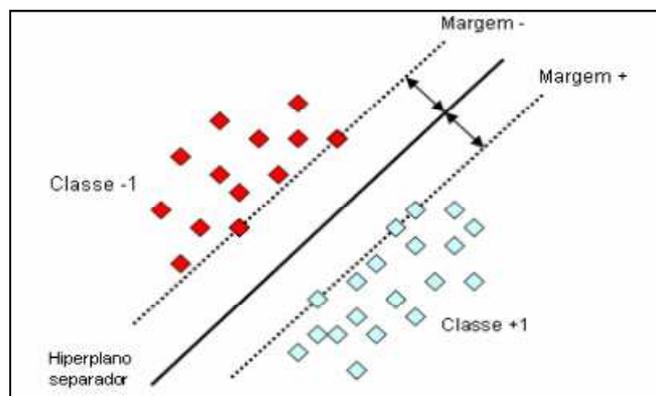


FIGURA 22 - Separação ótima de dois conjuntos linearmente separáveis

FONTE: Beltrami (2009)

Em linhas gerais, o SVM procura uma função (reta no caso R^2 , plano no caso R^3 e hiperplano no caso de R^n , com $n \geq 4$) que possua a mesma distância para os elementos de ambas as classes controlando, assim, a capacidade da função de decisão na busca (GÓES, 2012).

Nas figuras 21 e 22 anteriores apresentam exemplos em R^2 com conjuntos de dados linearmente separáveis. Assim, os vetores de suportes a e b na figura 23 seguinte são encontrados na fase de treinamento e definem a função de decisão ótima, sendo esta o classificador procurado.

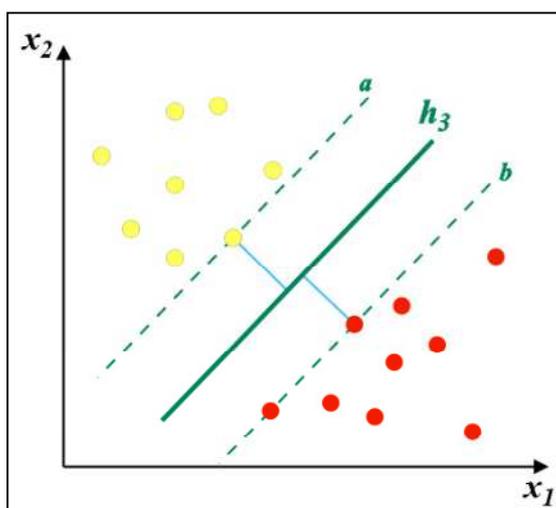


FIGURA 23 – Vetores suportes e função de decisão
FONTE: Góes (2012)

No caso de conjuntos em que os dados não são separáveis linearmente, o SVM utiliza a função Kernel (função presente no algoritmo que calcula a função de classificação) para projetar estes dados em um espaço em que seja possível separá-los linearmente, denominado de espaço de características com uma dimensão mais elevada (GÓES, 2012). Desta forma, pode ser resolvido pelo SVM, uma vez que a função de decisão é separável linearmente apenas no espaço de característica e não no espaço de entrada dos padrões conforme mostrado na figura 24 seguinte.

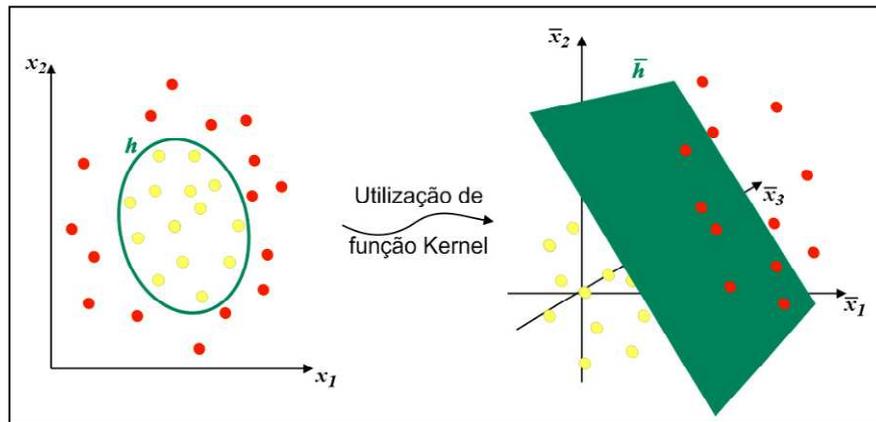


FIGURA 24 – Espaço de entrada dos padrões e espaço de características
 FONTE: Góes (2012)

Vapnik (1999) afirma ainda que para se ter um bom entendimento do SVM, pode-se comparar sua estrutura com a estrutura de uma RNA, sendo que nesta comparação a RNA possui apenas uma camada escondida, como mostra a figura 25.

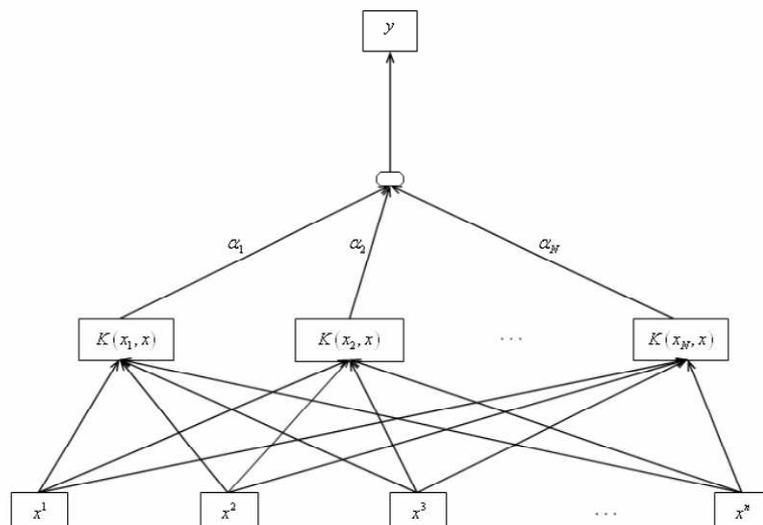


FIGURA 25 – Representação do SVM na estrutura de RNA
 FONTE: Vapnik (1999)

Desta forma, pode-se dizer que o funcionamento dos algoritmos SVM está baseado na metodologia de aprendizado supervisionado fazendo com que um sistema aprenda uma função a partir de uma amostra de dados.

4.4.2.1 APRENDIZADO SUPERVISIONADO DA SVM

No aprendizado supervisionado são utilizadas duas etapas: a de treinamento e a de teste. Na etapa de treinamento são fornecidos exemplos de entrada e saída, cujas coordenadas são conhecidas. O objetivo principal desta fase é ajustar os parâmetros livres do sistema e então, determinar uma função que relacione os respectivos valores. Para isso, penalizam-se os erros de estimação e valorizam-se acertos, de forma que o sistema retenha informações relevantes e adquira conhecimento (CARVALHO, 2005).

O algoritmo deve ter a capacidade de, após observar o comportamento de vários pares de entrada e saída, repetir este comportamento e gerar saídas próximas de y_i a partir de novas entradas de x_i , que não fizeram parte da fase de treinamento, porém com as mesmas características, conforme ilustrado na figura 26.

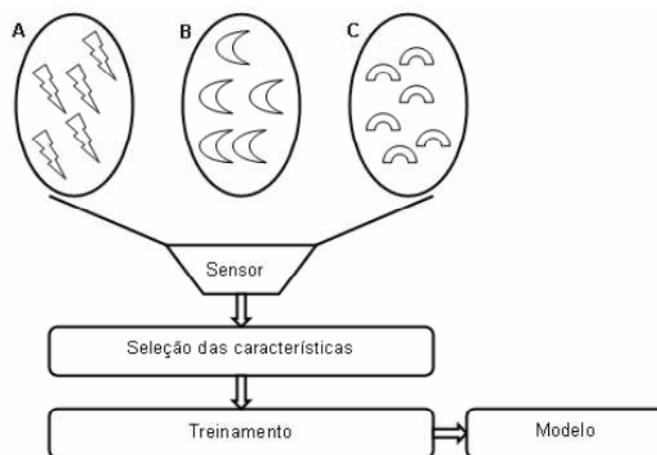


FIGURA 26 – Esquema de treinamento supervisionado

FONTE: Ales (2008)

Na fase de teste utiliza-se um conjunto de dados sobre o qual o método será efetivamente aplicado. Como são conhecidos apenas os valores de entrada dessa amostra, o intuito é utilizar a função determinada no treinamento para gerar as saídas desejadas e encontrar os resultados (BELTRAMI, 2009).

Neste contexto, a partir de um conjunto de vetores de treinamento pertencentes a duas classes linearmente separáveis, a tarefa do SVM é determinar uma função que classifique corretamente os dados de testes. Em outras palavras, significa encontrar um hiperplano separador com boa capacidade de generalização.

Com isso os novos dados de entrada são separados pelo hiperplano em suas classes já pré-definidas no treinamento, conforme ilustrado na figura 27.

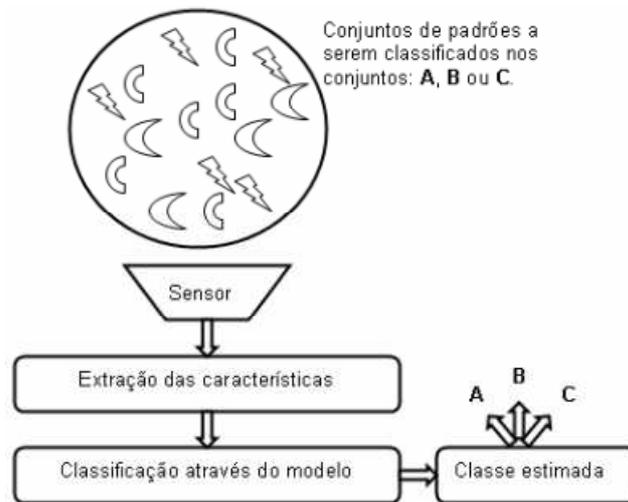


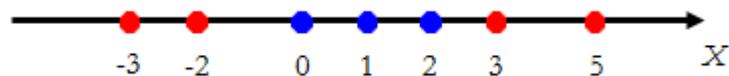
FIGURA 27 – Esquema de teste

FONTE: Ales (2008)

Para ocorrer sucesso (reconhecimento correto) é necessário que o conjunto de treinamento seja estatisticamente representativo, ou seja, represente os dados da problemática.

4.4.2.2 UM EXEMPLO DIDÁTICO DE CLASSIFICAÇÃO COM SVM

Seja o seguinte conjunto de exemplos de treinamento que não são linearmente separáveis:



Percebe-se que não há como realizar uma classificação destes pontos em classes; com isso, há a necessidade de elevar esses pontos para uma dimensão linearmente separável ($R^1 \rightarrow R^2$), conforme a tabela 3.

TABELA 3: Dados exemplo SVM

X_i	X_i^2	Classificação
-3	9	1
-2	4	1
0	0	-1
1	1	-1
2	4	-1
3	9	1
5	25	1

FONTE: Autor (2014).

Após esta etapa pode-se expressar graficamente os pontos e calcular o hiperplano ótimo que possa realizar a separação dos pontos em classes, conforme apresentado no gráfico 1 seguinte.

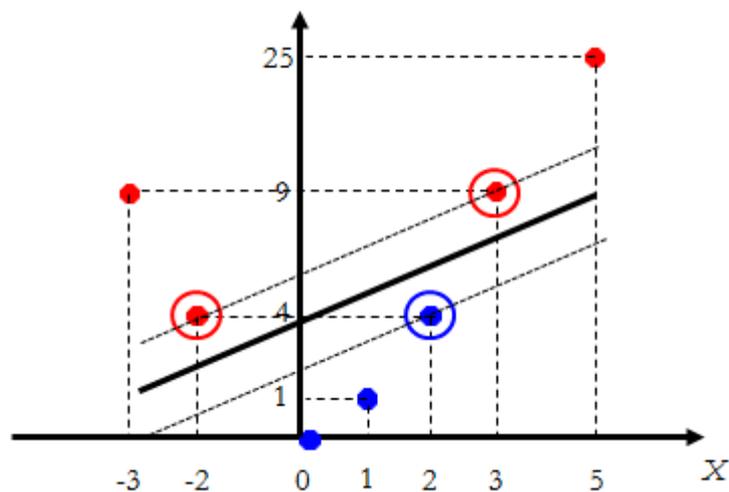


GRÁFICO 1 – Gráfico dos dados do exemplo SVM.

FONTE – Autor (2014)

Neste caso o hiperplano separador ótimo é dado pela equação 1 e a margem superior é dada pela equação 2, assim como a margem inferior é dada pela equação 3.

$$(W_0X + b_0 = 0) \quad (1)$$

$$(W_0X + b_0 = -1) \quad (2)$$

$$(W_0X + b_0 = 1) \quad (3)$$

Com isso, escolhem-se os pontos onde o hiperplano irá fazer separação das classes, neste exemplo os pontos (-2,4), (2,4) e (3,9), conforme destacado no gráfico 1 anterior.

Utilizando a equação 2 ($W_0X + b_0 = -1$) nos pontos (-2,4) e (3,9) e a equação 3 ($W_0X + b_0 = 1$) no ponto (2,4) tem-se:

$$\text{Equação 2: } W_1X_1 + W_2X_2 + b_0 = -1$$

$$-2w_1 + 4w_2 + b_0 = -1, \text{ com isso temos que, } b_0 = -1 + 2w_1 - 4w_2;$$

$$\text{Equação 3: } W_1X_1 + W_2X_2 + b_0 = 1$$

$$2w_1 + 4w_2 + b_0 = 1;$$

$$3w_1 + 9w_2 + b_0 = 1.$$

Fazendo as devidas substituições encontraremos os seguintes valores: ($W_1 = 1$; $W_2 = -1$ e $b_0 = 5$). Estes valores utilizados na equação 1 do hiperplano separador ótimo servirão para classificar os próximos pontos em +1 ou -1.

Caso adicionemos outros pontos no gráfico (4,16) e (3,2), por exemplo, então substitui-se na equação 2 com os valores calculados acima. Desta forma encontraremos:

$$W_0X + b_0 = 0$$

$$1 * 4 - 1 * 16 + 5 = -7 \text{ (classe - 1)}$$

$$1 * 3 - 1 * 2 + 5 = 6 \text{ (classe + 1)}$$

E assim repete-se este procedimento para os próximos ponto, como isso pode-se determinar a classe em que cada novo ponto pertencerá.

4.4.2.3 CONSIDERAÇÕES SOBRE APLICAÇÃO COMPUTACIONAL DO SVM

Na aplicação realizada à problemática desta Dissertação foi utilizado primeiramente o argumento *svmtrain*, do software *Matlab R2012b*, para o treinamento, com duas matrizes nos argumentos, conforme equação 4.

$$\text{Treino} = \text{svmtrain}(\text{Entrada_Treinamento}, \text{Saida_Treinamento}) \quad (4)$$

A matriz “*Entrada_Treinamento*” possui em seus vetores os valores de x_i e a matriz “*Saida_Treinamento*” possui um único vetor com os valores da faixa que cada padrão (linha da matriz “*Entrada_Treinamento*”) tem como resposta. Conforme mostrado nas matrizes 1 e 2.

$$Entrada_Treinamento = \begin{bmatrix} X_{11} & X_{21} & \dots & X_{61} \\ X_{12} & X_{22} & \dots & X_{62} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1n} & X_{2n} & \dots & X_{1n} \end{bmatrix} \quad (1)$$

$$Saida_Treinamento = \begin{bmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1n} \end{bmatrix} \quad (2)$$

Na sequência, foi utilizado o conjunto de teste, para realizar a classificação, aqui escrito em forma de matriz denominada “*Entrada_Testes*”, e o resultado de “Treino” com o argumento *svmclassify*, conforme a equação 5 com a finalidade de verificar o percentual de classificação corretas destes novos dados.

$$Classificação = svmclassify(Treino, Entrada_Testes) \quad (5)$$

A matriz *Entrada_Testes* tem o mesmo padrão da matriz *Entrada_Treinamento* porém com dados ainda não apresentados na fase de treinamento, ou seja, são dados totalmente novos.

5 APLICAÇÃO DAS TÉCNICAS DE RP

Este capítulo apresenta a obtenção do banco de dados, sua estruturação, os resultados obtidos com a aplicação das técnicas de RP, RNAs e SVM, tanto para a previsão quanto para a classificação de padrões.

5.1 OBTENÇÃO E ESTRUTURAÇÃO DO BANCO DE DADOS

Para uma análise mais precisa do comportamento do teor cáustico na lama vermelha foi elaborado um gráfico de acompanhamento. Este gráfico serve para verificar a variabilidade do processo, que se mostra aleatória e inconstante. Tal gráfico também poderá auxiliar a equipe de processo em análises prévias para identificar o quanto estes dados são influenciados por ações individuais adotadas por quem opera o processo.

O gráfico 2 representa os resultados das amostras de percentuais de soda na lama em g/l em um determinado período de tempo. Pode-se observar que os resultados possuem grande variabilidade, ou seja, não está sob controle, e sendo assim, a empresa não é capaz de atender às especificações exigidas. É importante salientar que as especificações são o que se espera atingir e o limite de controle é o que o processo pode fornecer como consistência.

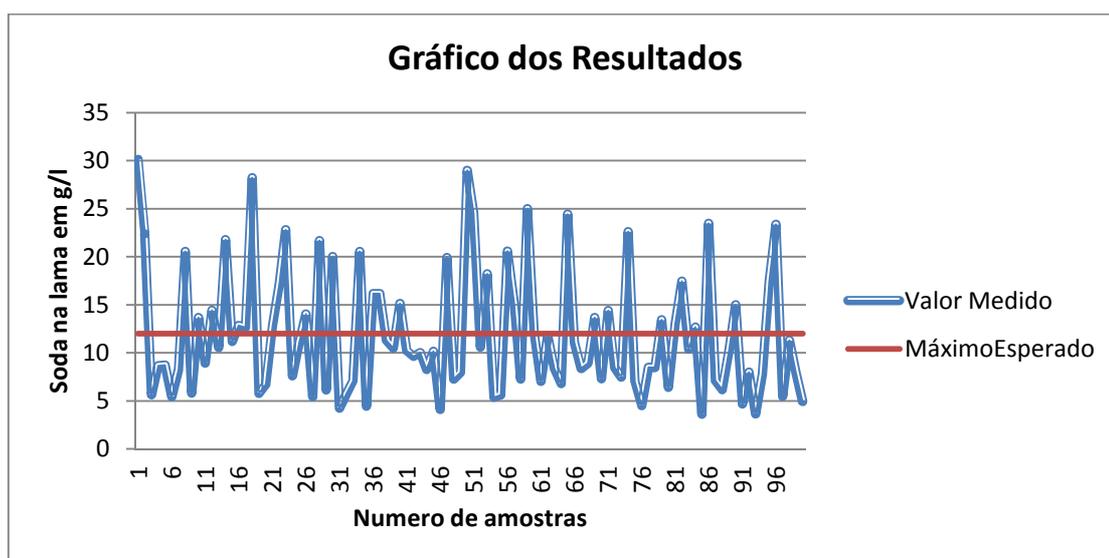


GRÁFICO 2 – Gráfico dos Resultados das Análises

FONTE – Autor (2014)

Como já comentado anteriormente, o objetivo deste estudo é atingir as especificações definidas pela empresa, minimizando o teor de soda cáustica contida no rejeito do processo, tendo em sua metodologia base o processo Descoberta de Conhecimento em Bases de Dados - *Knowledge Discovery In Data Bases (KDD)*.

Para tanto, a execução deste trabalho iniciou-se pela coleta de dados aqui apresentada, que “alimentarão” as técnicas de RP. Os resultados desta aplicação poderão ser então utilizados para a construção de um modelo matemático visando a minimização dos teores de soda cáustica do rejeito, assim como dos custos de todo o processo para tal fim.

O procedimento de amostragem da pesquisa foi realizado através de observação sistemática, pois conforme Silva e Menezes (2005), este tipo de coleta de dados tem planejamento e realiza-se em condições controladas de forma que os resultados sejam de acordo com o pré-estabelecido.

As amostras para análises desta pesquisa foram coletadas e organizadas em uma planilha eletrônica, gerando assim a base de dados para o problema em questão. Na sequência, uma primeira análise foi realizada pelo especialista de processo da referida área de lavagem de lama, que os identificou como sendo os atributos de entrada e a variável de saída do referido processo, excluindo do banco de dados registros não consistentes, ou seja, dentro do processo KDD esta análise é considerada como a limpeza dos dados.

Cada grupo de dados coletados para este estudo obedeceu todos os critérios de coleta de amostra da empresa, a fim de evitar falhas ou erros de leitura nos instrumentos. Essas amostras foram coletadas diariamente, sendo quatro amostragens por dia, em horários pré-estabelecidos pelo especialista da área, perfazendo um total de 800 amostras.

A tabela 4 seguinte, apresenta alguns desses registros com os seguintes atributos: Densidade da Polpa em g/cm^3 (X_1), Vácuo em kgf (X_2), Nível da bacia do filtro rotativo em percentagem (X_3), Rotação do filtro de lavagem em rotações por minuto (X_4), Condensado de lavagem da lama ou Spray, em m^3/h (X_5) e Teor Total de Soda em gramas por litro, g/l (X_6) e como uma variável de saída, a Concentração cáustica no rejeito da lama vermelha, também em g/l (Y).

TABELA 4 – Dados do processo de lavagem de lama.

Amostras	Densidade (g/cm ³)	Vácuo (kgf)	Nível de Bacia (%)	Rotação (rpm)	Spray (m ³ /h)	TTS (g/l)	Concentração (g/l)
N	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	Y
1	1,49	0,51	25,55	1,60	24,81	66,57	5,85
2	1,52	0,51	25	1,60	25,23	70,43	6
3	1,51	0,52	19,64	1,60	22,25	68,4	14,65
4	1,53	0,51	19,43	1,60	23,79	64,77	16,3
5	1,56	0,52	24,3	2,00	30	66,28	14,25
6	1,53	0,51	16,84	1,70	29,01	67,88	14,65
7	1,56	0,5	20,65	1,80	29,53	66,06	11,05
8	1,52	0,51	19,45	1,80	30,27	69,17	8,45
9	1,51	0,5	28,47	1,80	28,86	62,35	11,2
10	1,53	0,51	24,94	1,70	26,56	61,99	10,52
11	1,55	0,51	16,47	1,70	26,97	82,56	6,8
12	1,53	0,5	26,52	1,90	27,04	76,15	7,8
.
.
.
790	1,5	0,33	20,97	1,70	22,3	48,85	16,55
791	1,54	0,37	18,27	1,80	21,64	53,81	17,25
792	1,54	0,33	21,71	1,80	22,26	48,84	17,95
793	1,5	0,33	18,12	2,20	22,99	56,31	18,56
794	1,51	0,35	22,24	1,40	22,15	59,57	14,53
795	1,43	0,35	20,45	2,00	21,9	60,42	13
796	1,43	0,4	19,89	2,00	21,99	64,32	11,48
797	1,43	0,43	22,59	2,10	23,74	62,7	9,95
798	1,43	0,45	18,66	2,20	23,71	64,1	8,49
799	1,43	0,44	19,96	2,10	24,02	58,83	8,16
800	1,52	0,42	18,16	2,10	23,76	59,99	8,24

FONTE – Autor (2014)

Nesta base de dados foram necessárias transformações nos registros para que pudessem ser aplicadas as técnicas de RP, sendo a primeira a separação dos 800 registros em duas classes definidas pelo especialista da área, que se baseou nos resultados de laboratório que analisa o teor de soda contida na lama vermelha, ou seja, os resultados da variável concentração (g/l) que equivale a Y (saída/resultado).

Atualmente, é considerado como aceitável o valor de soda na lama de no máximo 12 g/l, ou seja, valores menores ou iguais ao estipulado são considerados resultados de boa qualidade (bons), já valores acima deste são considerados resultados de baixa qualidade (ruins), pois são considerados com alto teor de soda

cáustica, o que além de elevar os custos operacionais aumentam os riscos de contaminação ambiental, já que esta lama fica depositada em bacias de contenção a céu aberto.

Sendo assim, o banco de dados foi dividido em dois grupos, o considerado “bom” e o considerado “ruim” pelos especialistas de processo da área, ficando assim definida uma saída binária (“0” ou “1”), sendo considerado como “0” os resultados bons, ou seja, iguais ou menores a 12 g/l, e considerado como “1” os resultados ruins, maiores do que 12 g/l. Após esta classificação preliminar, observou-se que 60% de amostras ficaram classificadas como boas e 40% como ruins. A tabela 5 resume todas estas informações.

TABELA 5: Pré-classificação binária das amostras

Pré-classificação das amostras			
Resultados	Bons (≤ 12 g/l)	Ruins (> 12 g/l)	Total
Amostras	480	320	800
Classificação	“0”	“1”	
Percentual	0,6	0,4	

FONTE: Autor (2014)

Foram iniciados os procedimentos de RP, utilizando-se as técnicas de RNA e SVM, utilizando-se para o treinamento das técnicas todos os padrões, aqui denominada, 1ª. metodologia, porém os resultados não foram satisfatórios. Vários testes foram realizados com ambas as técnicas, porém o percentual de acerto ficou extremamente baixo, não atingindo 60%. Tais resultados serão melhor explicitados na seção 5.2 onde se mostram todos os resultados desta pesquisa.

Após os resultados insatisfatórios citados, decidiu-se aplicar outra metodologia para estruturação dos dados aplicada em Góes (2012).

Esta nova metodologia, aqui denominada de 2ª. metodologia, consiste em dividir o total da amostra, neste caso $n = 800$, em outros grupos, considerados subgrupos, de menor tamanho com mesmo número de elementos, sendo que para este estudo cada subgrupo ficou com 160 observações, ou seja, $n = 160$. Tal divisão dos subgrupos foi em função dos dados de saída em (Y) ordenados em ordem crescente. Com a subdivisão obteve-se 5 faixas de classificação para o valor esperado na variável de saída (Y), classificadas como Excelente, Muito bom, Bom, Ruim e Muito ruim. Tal classificação foi adotada em função dos valores esperados

em (Y), respeitando os limites estabelecidos pelos especialistas do referido processo.

A tabela 6 mostra a criação dos subgrupos, bem como a quantidade de elementos de cada subgrupo, seus limites e sua classificação.

TABELA 6: Criação dos subgrupos

Subgrupos	Qtde amostras			Limites		
	Início	Fim	% participação	Inferior	Superior	Classificação
A	1	160	0,20	2,2	6,8	Excelente
B	161	320	0,20	6,9	8,9	Muito bom
C	321	480	0,20	9	12	Bom
D	481	640	0,20	12,1	17,2	Ruim
E	641	800	0,20	17,3	41	Muito ruim

FONTE: Autor (2014)

Na aplicação das técnicas de RP fez-se a divisão do problema em subproblemas da seguinte maneira: ao invés de se ter a saída, por exemplo, $A = 0$, com 60% dos dados e $B = 1$, com 40% dos dados, é realizado nesta segunda metodologia um treinamento com os dados do subconjunto $A = 0$ (20% dos dados) e um conjunto $K^* = 1$ composto dos demais registros do treinamento (B, C, D e E), ou seja, 80% dos dados. Em seguida realiza-se outro treinamento com os registros de $B = 0$ e um conjunto $K^* = 1$ composto dos registros de treinamento de (C, D e E). Em seguida, o treinamento como os registros do subgrupo $C = 0$ e um conjunto $K^* = 1$, composto por (D e E). E finalmente, um conjunto com os registros dos dados do subgrupo $D = 0$ e o conjunto $K^* = 1$ com os elementos dos subconjuntos (E). Desta maneira todos os dados passam por treinamento, sendo esta técnica conhecida como “um contra todos” (GOES, 2012).

O gráfico 3 mostra como os dados foram divididos, em função da variável Y, a fim de viabilizar esta técnica.

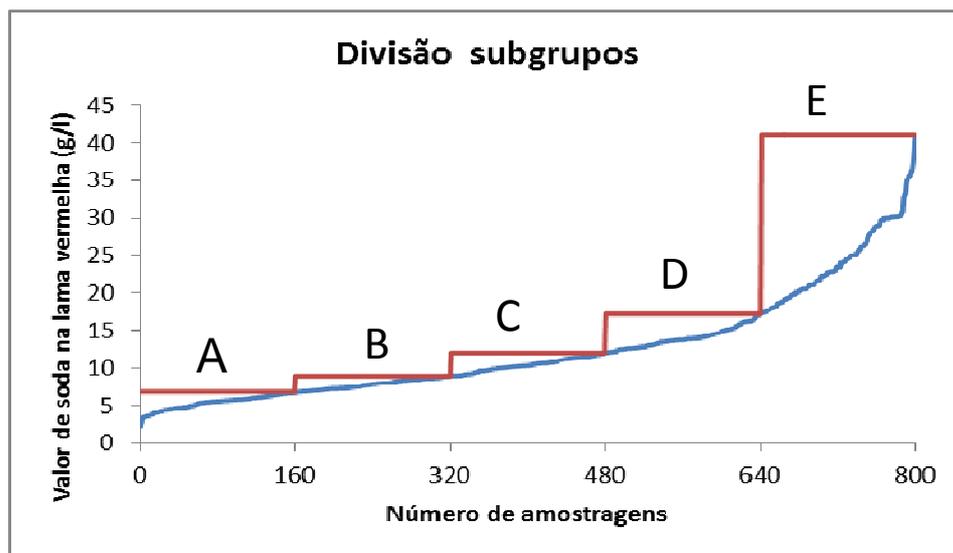


GRÁFICO 3: Gráfico subgrupos

FONTE: Autor (2014)

Nos testes foi utilizado a Validação Cruzada (*Cross-Validation*), tanto para a técnica SVM, quanto para a RNA. Esta forma de validação dos resultados foi escolhida, pois, trata-se de uma avaliação completa e comumente utilizada para estimar a taxa de acerto de classificadores (GRANATYR, 2011).

O princípio da técnica de validação cruzada é simples. Aqui, a base de dados T com n instâncias é subdividida em f amostras geradas aleatoriamente. O tamanho de cada amostra é igual ao número de instâncias n dividido pelo número de amostras f . Tal fração pode gerar um valor aproximado para o caso n de uma divisão não exata (KOHAVI, 1995).

O método de validação cruzada utilizada neste estudo foi do tipo *k-fold*, que consiste em dividir o conjunto total de dados em k subconjuntos mutuamente exclusivos do mesmo tamanho e, a partir disto, $k-1$ subconjuntos são utilizados para treinar a técnicas e o conjunto restante é utilizado para testá-la. Repete-se este procedimento por k vezes podendo-se, desta forma, calcular a acurácia do modelo. A figura 28 mostra o esquema realizado pelo *k-fold*.

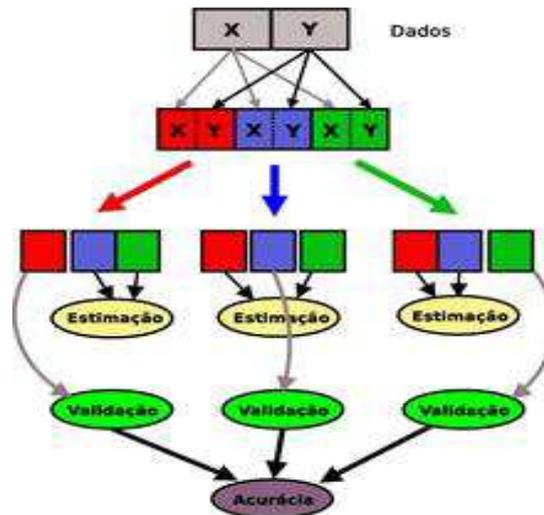


FIGURA 28: Método *K-folds* ($k=3$)

FONTE: Autor (2014)

Ao final das k avaliações calculam-se a acurácia levando em consideração os erros encontrados.

5.2 RESULTADOS OBTIDOS NA CLASSIFICAÇÃO

Nesta seção apresentam-se os resultados obtidos primeiramente com as técnicas utilizadas, SVM e RNAs, na classificação dos padrões.

5.2.1 RESULTADOS COM SVM

Inicialmente os testes com o algoritmo do SVM foram realizados com o banco de dados estruturado em duas classes, “Bom” e “Ruim”, conforme descrito na seção 5.1 anterior.

Para esta técnica utilizaram-se cinco tipos diferentes de funções de *Kernel*, a fim de maximizar a taxa de acerto. Tais funções estão disponíveis no *Software Matlab R2012b* e são assim denominadas:

- a) *linear - linear kernel*;
- b) *quadratic - quadratic kernel*;
- c) *polynomial - polynomial kernel*;
- d) *rbf - gaussian radial basis function kernel*;
- e) *mlp - multilayer perceptron kernel*.

Foi utilizado o procedimento de treinamento e teste da seguinte forma: utilizou-se 70% do total de dados para treinamento da rede e 30% para a validação, mantendo a proporcionalidade dos dados considerados bons e ruins nas parcelas de treinamento e validação. O modelo do treinamento do modelo SVM dá-se pela equação 6.

$$\text{Treino} = \text{svmtrain}(\text{EntradaTreino}, \text{SaidaTreino}, \text{'Kernel_Function'}, \text{'rbf'}) \quad (6)$$

Já o teste com os demais 30% é realizado através da equação (7).

$$\text{Classificação} = \text{svmclassify}(\text{Treino}, \text{EntradaValidacao}) \quad (7)$$

Este procedimento foi repetido por 3.000 vezes com cada uma das funções *kernel* citadas anteriormente. A máxima acurácia obtida foi de 57%, considerada insuficiente para as necessidades do processo pesquisado. Os resultados obtidos são apresentados na tabela 7.

TABELA 7: Classificação com SVM

% acertos para n = 800					
Iterações	Funções do SVM				
	Linear	Rbf	mlp	Quadratic	polynomial
150	0,55	0,48	0,55	0,56	0,59
200	0,52	0,5	0,45	0,53	0,58
250	0,51	0,47	0,47	0,5	0,55
Média	0,53	0,48	0,49	0,53	0,57
Desvpad	0,02	0,02	0,05	0,03	0,02

FONTE: Autor (2014)

Aplicou-se, então, a 2ª metodologia na divisão das classes para aplicação da referida técnica SVM. Após a devida estruturação dos dados foram realizados testes com as funções *kernel* novamente, para verificar qual delas apresenta uma melhor aderência aos dados do problema, proporcionando uma melhor taxa de acertos. Com esta nova estruturação dos dados e repetidos os testes com o algoritmo apenas a função Gaussiana (rbf) apresentou boas taxas de acerto, conforme apresentado na tabela 8.

TABELA 8: Resumo dos resultados com SVM

n = 800 em 5 subgrupos					
% de acertos com 3000 iterações					
Funções	Linear	Rbf	mlp	quadratic	polynomial
A+BCDE	0,51	0,79	0,47	0,24	0,21
B+CDE	0,56	0,77	0,48	0,58	0,25
C+DE	0,56	0,79	0,46	0,58	0,34
D+E	0,59	0,79	0,44	0,62	0,50
média	0,55	0,79	0,46	0,51	0,33

FONTE: Autor (2014)

Os resultados apresentados na tabela 8 anterior, mostram que se obteve uma melhora nos resultados com esta 2ª metodologia quando comparados com a 1ª metodologia (tabela 7). Com isso obteve-se uma elevação na taxa de acertos de 53% para 79%. Trata-se de uma taxa de acerto considerada relativamente boa pelos especialistas da área da referida problemática.

5.2.2 RESULTADOS COM AS RNAs

Os dados também foram submetidos à técnica de RNA mantendo a mesma estruturação. Inicialmente criou-se cinco RNAs, a fim de verificar a aderência dos dados à referida técnica. Cada uma das redes ficou com seis neurônios na camada de entrada, um número pré-determinado de neurônios na camada escondida (5; 10; 15; 20 e 25) e apenas um neurônio na camada de saída. Tais redes utilizaram 10.000 iterações (épocas) como critério de parada. Na tabela 9 mostra-se o resumo dos resultados obtidos com a técnica de RNA.

TABELA 9 - Resumo dos resultados com RNA

n = 800 em 5 subgrupos					
% de acertos com 10.000 iterações					
Escondidos	5	10	15	20	25
A+BCDE	0,71	0,83	0,82	0,85	0,85
B+CDE	0,69	0,70	0,72	0,77	0,88
C+DE	0,70	0,81	0,73	0,76	0,83
D+E	0,71	0,81	0,82	0,83	0,88
Média	0,70	0,79	0,77	0,80	0,86

FONTE: Autor (2014)

Vale ressaltar que os testes realizados com as RNAs cumpriram os mesmos procedimentos e estruturação dos dados da 2ª. metodologia que foram utilizados para os testes com o SVM, conforme explicitado na seção 5.1, a fim de evitar distorções na comparação dos resultados. Ou seja, os dados foram submetidos nas mesmas condições de treinamento e testes, com a mesma metodologia da validação cruzada.

Observa-se na tabela 9 anterior que os resultados das taxas de acertos com as RNAs apresentam-se melhores, mais satisfatórios para a problemática deste trabalho comparativamente aos resultados provenientes do SVM. Observa-se também que quando se compara as RNAs entre elas com diferentes neurônios na camada escondida, a rede com 25 neurônios apresentou melhores resultados, sendo 86% no percentual de acertos.

A partir destes resultados pode-se inferir que a técnica do SVM para classificação não apresentou resultados tão bons quanto os obtidos pela RNAs. Sendo assim, optou-se refinar os resultados da RNAs, conforme explicado a seguir.

Após estes testes iniciais, novas RNAs foram criadas com objetivo de identificar melhorias nas taxas de acertos dos padrões reconhecidos. Os novos testes adicionais realizados foram efetuados variando o número de neurônios na camada escondida, sendo de 1 até 25 neurônios, realizando o incremento de um neurônio de cada vez. Nestes testes, a taxa de aprendizagem da rede variou de 0,1 até 0,9 também utilizando o incremento de 0,1 por vez para cada neurônio. Na tabela 10 mostra-se o resumo de como os testes foram processados e tabulados.

TABELA 10: Estruturação dos resultados dos testes

Taxa Aprendizagem	Nº Neurônios na Camada Escondida								
	1	2	3	.	.	.	23	24	25
0,1									
0,2									
0,3									
0,4									
0,5									
0,6									
0,7									
0,8									
0,9									

Taxa de acerto na Classificação de Padrões

FONTE: Autor (2014)

Na arquitetura das redes criadas foi utilizado o algoritmo de aprendizagem *backpropagation*. Sua composição sempre feita com os 6 neurônios na camada de entrada, um número variável de neurônios, de 1 a 25, na camada escondida e um neurônio na camada de saída, representando a variável de saída da rede.

Dentre as funções utilizadas na configuração da rede, estão a *logsig* de ativação e *traingdx* para treinamento. Tendo como critério de parada do treino, o limite de 10.000 épocas (iterações). Na tabela 11 mostra-se a configuração das RNA's criadas.

TABELA 11: Parâmetros de treinamento da RNA para a classificação de padrões.

Parâmetros	Valor
Taxa de aprendizagem	0,1 a 0,9
Neurônios na camada oculta	1 e 25
Função de transferência da camada oculta	Logística (<i>logsig</i>)
Função de transferência da camada de saída	Logística (<i>logsig</i>)

FONTE: Autor (2014)

A partir dessa etapa seguiram-se as fases de treinamento e teste da rede. As taxas de acertos para cada rede, levando em consideração a variação tanto com o número de neurônios na camada escondida quanto às taxas de aprendizagem são apresentadas na tabela 12. A tabela completa com todos os resultados encontra-se no Apêndice A deste trabalho.

TABELA 12: Melhores taxas de acertos das RNA's para a classificação de padrões.

Taxa Aprendizagem	Nº Neurônios na Camada Escondida												
	9	10	11	12	13	14	15	16	17	18	19	20	23
0,1													0,76
0,2				0,79									
0,3					0,80								
0,4					0,93								
0,5	0,88												
0,6								0,91					
0,7										0,91			
0,8												0,91	
0,9					0,84								

FONTE: Autor (2014)

Após os vários testes realizados, pode-se observar que a técnica de RNA apresentou resultado bastante satisfatório para a classificação de padrões. Como mostrado na tabela 12 anterior, conseguiu-se modelar uma rede com uma taxa de acertos de 93%, ficando com a seguinte configuração: Seis neurônios na camada de entrada, 13 neurônios na camada oculta/intermediária e um neurônio na camada de saída (6 – 13 – 1), sendo que foi utilizado 10.000 iterações como critério de parada e uma taxa de aprendizagem de 0,4.

Além das taxas de acertos medias para cada RNA criada, foi calculado também seu erro médio quadrático (MSE), afim de ratificar qual rede apresenta melhores resultados de classificação. O MSE consiste na diferença entre os valores reais de soda cáustica contidos na lama vermelha (V_r ou “Concentração” da tabela 4) e os valores estimados pela RNA (V_e) com o banco de dados dos 30% reservados à validação da rede. O MSE foi calculado de acordo com a equação 8 onde n é o número de padrões utilizados ($n = 560$ para o treinamento e $n = 240$ para o teste).

$$MSE = \frac{\sum(V_r - V_e)^2}{n} \quad (8)$$

Na tabela 13 mostram-se como os dados foram tabulados em relação ao MSE atingido para cada neurônio inserido e a cada incremento na taxa de aprendizagem da rede.

TABELA 13: Estruturação dos resultados dos MSE's das RNA's de Classificação

Taxa Aprendizagem	Nº Neurônios na Camada Escondida								
	1	2	3	.	.	.	23	24	25
0.1									
0.2									
0.3									
0.4									
0.5						MSE atingido			
0.6									
0,7									
0,8									
0,9									

FONTE: Autor (2014)

Após esta estruturação seguiram-se as fases de treinamento e teste da rede, obtendo-se o MSE para cada rede com cada variação, tanto com o número de neurônios escondidos, quanto em relação às taxas de aprendizagem. Na tabela 14 mostram-se os resultados com os menores MSE's encontrados. A tabela completa com todos os resultados encontra-se no Apêndice B deste trabalho.

TABELA 14: Resultados dos MSE's das RNA's de Classificação

Taxa Aprendizagem	Nº Neurônios na Camada Escondida												
	9	10	11	12	13	14	15	16	17	18	19	20	23
0,1													0,060
0,2				0,045									
0,3					0,042								
0,4					0,005								
0,5	0,018												
0,6								0,010					
0,7										0,007			
0,8												0,009	
0,9					0,036								

FONTE: Autor (2014)

Com os resultados apresentados na tabela 14 anterior, pode-se ratificar que a técnica de RNA apresenta bons resultados RP, tal inferência confirma-se pelo baixo MSE atingido na rede com 13 neurônios na camada intermediária e 0,4 de taxa de aprendizagem.

5.3 RESULTADOS OBTIDOS NA PREVISÃO

Nesta seção apresentam-se os resultados obtidos com a técnica de RNAs, utilizada para a previsão.

5.3.1 RESULTADOS COM RNAs

Os dados foram submetidos à técnica de RNA, porém nesta seção, a abordagem desta técnica apresenta-se como previsão dos resultados em y a partir do treinamento dos dados de entrada x_i . Da mesma forma como na classificação, foram inicialmente criadas RNAs, a fim de verificar a aderência dos dados à referida técnica, sendo que cada rede possui seis neurônios na camada de entrada, um

número pré-determinado de neurônios na camada escondida e apenas um neurônio na camada de saída. Tais redes utilizaram 10.000 iterações (épocas) como critério de parada.

Nas RNAs criadas, o número de neurônios na camada escondida foi sendo variado, sendo de 1 até 25 neurônios, realizando o incremento de um neurônio por vez. Nestas RNAs também foram sendo alteradas a taxa de aprendizagem da rede, variando os valores de 0,1 até 0,9 também utilizando o incremento de 0,1 por vez para cada neurônio, desta forma foram criadas 225 RNA's.

Da mesma forma que foi realizada na Classificação, também nesta etapa de previsão, para cada RNA criada foi calculado o MSE após o seu treinamento, afim de verificar qual rede apresenta melhores resultados de previsão. O MSE calculado obedece a mesma estrutura da equação 8 anterior.

A tabulação dos MSE's neste processo de previsão, também mantém a mesma estruturação mostrada na tabela 13 anterior.

Após a estruturação seguiram-se as fases de treinamento e teste da rede, obtendo-se o MSE para cada rede com cada variação, tanto com o número de neurônios escondidos, quanto em relação às taxas de aprendizagem. Na tabela 15 mostram-se os resultados com os menores MSE's encontrados. A tabela completa com todos os resultados encontra-se no Apêndice C deste trabalho.

TABELA 15: Resultados dos MSE's das RNA's de Previsão

Taxa Aprendizagem	Nº Neurônios na Camada Escondida												
	1	2	3	.	.	.	19	20	21	22	23	24	25
0,1													0,0303
0,2							0,0294						
0,3									0,0167				
0,4												0,0151	
0,5											0,0021		
0,6												0,0162	
0,7												0,0016	
0,8							0,0156						
0,9											0,0153		

FONTE: Autor (2014)

Com todos os MSE's medidos, pode-se observar que quanto maior a taxa de aprendizagem e maior o número de neurônios na camada escondida, o erro da rede

fica cada vez menor, porém em um determinado ponto a erro começa a elevar-se novamente. Tal fenômeno pode ser melhor visualizado na gráfico 4.

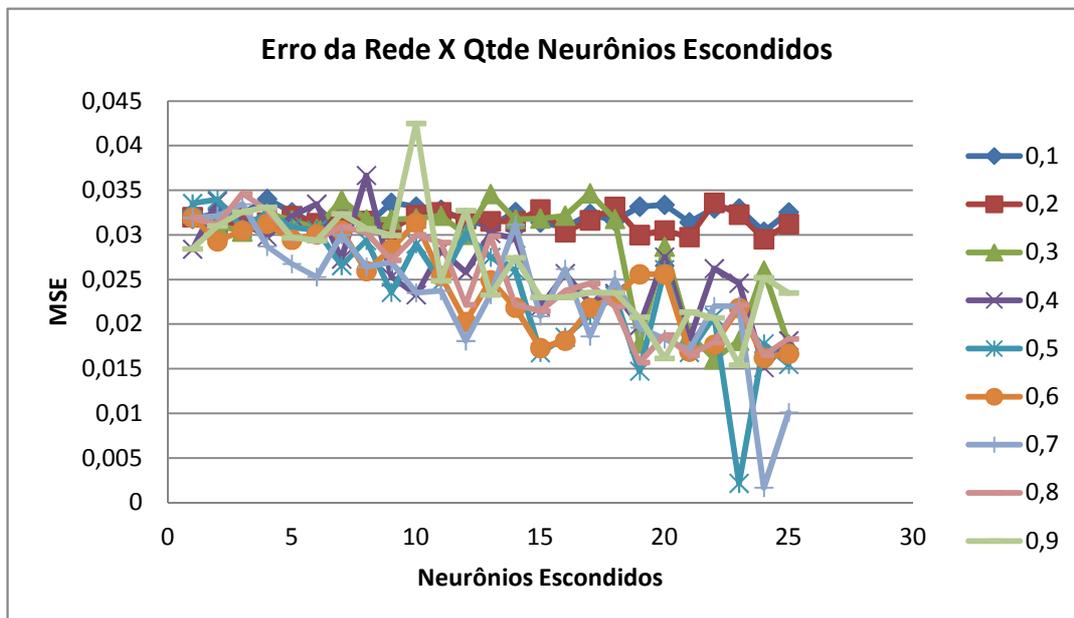


GRÁFICO 4 – Gráfico MSE's das RNA's criadas
 FONTE: Autor (2014)

É possível observar tanto na tabela 15 quanto no gráfico 4 anteriores que das 225 RNA's criadas, duas delas apresentaram um MSE consideravelmente baixo para a problemática em questão. Todas as redes criadas utilizaram o algoritmo de aprendizado *backpropagation*. Sua composição sempre feita com os 6 neurônios na camada de entrada e na camada escondida ou camada intermediária, compostas pelos números variados de neurônios (1 à 25), já a camada de saída é composta por apenas um neurônio representando a variável de saída da rede.

Dentre as funções utilizadas na configuração da rede, estão a *logsig* de ativação e *traingdx* para treinamento. Tendo como critério de parada do treino, o limite de 10.000 épocas (iterações). Na tabela 16 mostra-se a configuração das RNA's criadas.

TABELA 16: Parâmetros de treinamento para da RNA como Previsão

Parâmetros	Valor
Taxa de aprendizagem	0,1 à 0,9
Neurônios na camada oculta	1 e 25
Função de transferência da camada oculta	Logística (<i>logsig</i>)
Função de transferência da camada de saída	Logística (<i>logsig</i>)

FONTE: Autor (2014)

Podemos observar que a estrutura de uma RNA para previsão, segue os mesmos princípios da RNA para classificação, tanto no algoritmo de treinamento quanto nas funções de ativação da camada oculta e camada de saída. Porém o que difere uma RNA de previsão para uma RNA de classificação, está forma de como os resultados esperados em y são mostrados para a rede. Na classificação os resultados são esperados de forma binária (“0” ou “1”), enquanto na RNA para previsão, os dados são estruturados e esperados de forma contínua, podendo assumir qualquer valor no intervalo de “0” e “1”.

As duas RNA's de previsão criadas que apresentaram menores MSE's possuem as seguintes características conforme a tabela 17.

TABELA 17: Melhores RNA's desenvolvidas

RNA's Com menores MSE's					
Nº RNA	Camada de Entrada	Camada Oculta	Camada de Saída	Taxa de aprendizagem	MSE
123	6	23	1	0,5	0,002157
174	6	24	1	0,7	0,00167

FONTE: Autor (2014)

Com as RNA's citadas na tabela 17 anterior foram criados seus respectivos gráficos de previsão, conforme os gráficos 5 e 6.

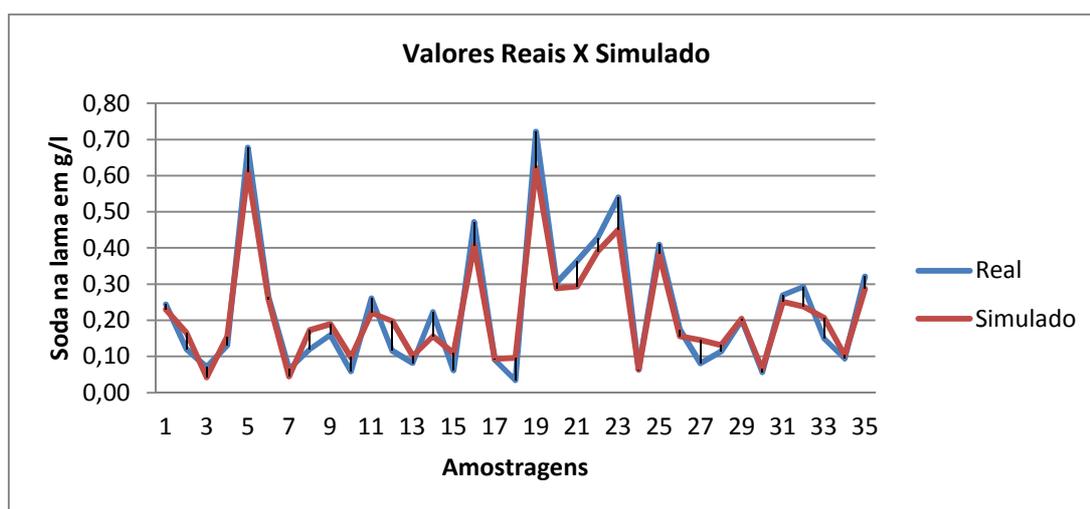


GRÁFICO 5 – Gráfico RNA nº 123

FONTE: Autor (2014)

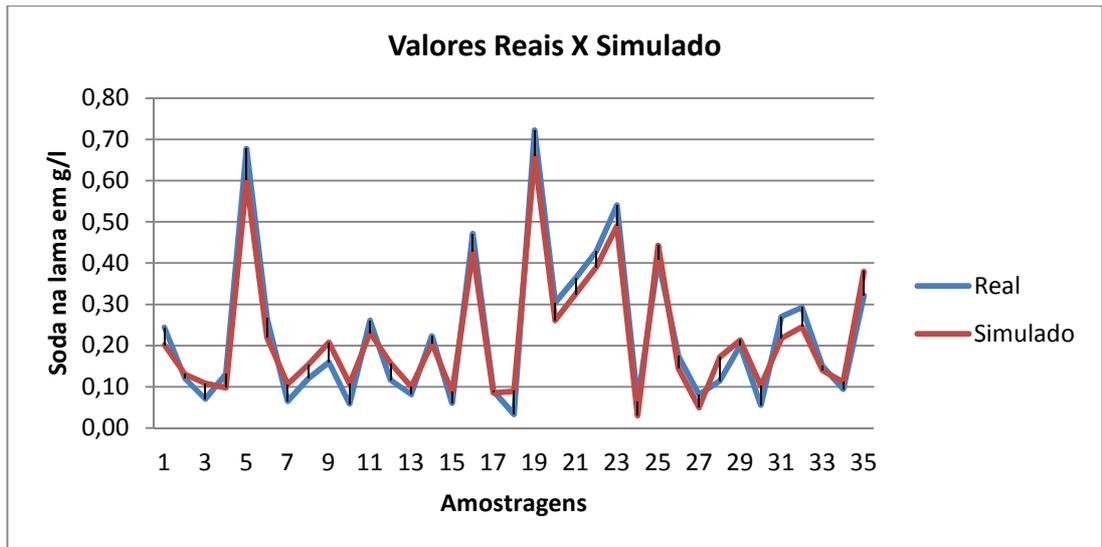


GRÁFICO 6 – Gráfico RNA nº 174

FONTE: Autor (2014)

Após os vários testes e RNA's criadas observa-se a existência uma boa aderência entre os valores reais e os valore previsto pela RNA, pois o comportamento dos valores estimados retrata o comportamento similar aos valores reais. Nota-se que existe o acompanhamento das tendências de subida e descida dos valores reais pelos valores estimados pelo modelo de previsão das RNA's.

6 CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS

O presente trabalho foi idealizado objetivando a utilização de técnicas de RP (RNA e SVM), de forma comparativa, a fim de identificar no rejeito do processo de produção de alumina, os padrões das variáveis de entrada que controlam o processo e geram os variados teores de soda cáustica neste rejeito, impossibilitado este de ser utilizado em outros processos devido sua grande variabilidade nos resultados. Além disso, o RP foi proposto também para realizar a previsão deste resultado, auxiliando a equipe de processo na tomada de ações preventivas. O presente estudo também poderá servir de base para a criação de um modelo matemático que possa auxiliar na minimização dos resultados de soda cáustica no rejeito do processo. Poderão ser determinados para quais valores das variáveis de entrada, os resultados desejados serão adequados para a empresa, com base no método de RP com melhor desempenho.

Foram Identificadas as variáveis que interferem no controle do teor cáustico do rejeito do processo e que foram utilizadas como variáveis de entrada do procedimento de classificação e previsão deste trabalho.

Os resultados da pesquisa mostram que para tal problemática a técnica de RNA obteve melhores resultados que o SVM, quando estamos procedendo a classificação, sendo possível identificar quais valores para as variáveis de entrada resultam nos teores de soda cáustica no rejeito, considerados bons pela empresa, ou seja, resultados que apresentem no máximo 12 g/l. A técnica de RNA obteve uma taxa máxima de acertos de 93%, o que representa um MSE de 0,005, enquanto que a taxa máxima de acertos para o SVM ficou em 79%, uma diferença significativa para a problemática.

A aplicação da técnica de RNA para a previsão se mostrou com elevada eficiência, com um MSE consideravelmente baixo (0,00167) e com isso, será possível informar aos especialistas de processo os possíveis problemas no sistema e necessidade de ações corretivas.

O modelo de previsão proposto apresenta melhores resultados em comparado com o método utilizado pela empresa, pois a empresa não trabalha com modelo de previsão, mas sim com gráficos de controle dos resultados de saída em y e, com isso, as ações são tomadas apenas de forma reativa após o resultado emitido pelo laboratório. Assim, as ações além de tardias nem sempre surtem efeitos positivos.

Com o RP realizado, pode-se verificar os limites inferiores e superiores de cada variável de controle, podendo auxiliar os especialistas da área a melhor controlá-las, podendo variar esses valores, dentro dos limites reconhecidos como padrões para bons resultados, evitando custos desnecessários. Com isso pode-se melhorar o sistema de lavagem.

Contudo deixa-se claro que a melhoria no sistema de controle, com previsão e classificação adequados dos padrões, por si só, não será a panacéia que solucionará de vez por todos os males do descontrole que levam aos resultados indesejáveis. Uma gestão eficiente do processo com a participação dos operadores e especialistas também é necessária, pois deverá estar alinhada com a estratégia da organização. Todo este procedimento visa não somente a redução do teor cáustico na lama vermelha, mas também os benefícios ao meio ambiente com o reaproveitamento deste material em outros processos produtivos.

Como propostas de estudos futuros tem-se a pretensão de utilizar os resultados da técnica de RP que apresentou o melhor desempenho, ou seja, RNA para previsão, para o desenvolvimento de um modelo matemático que seja capaz de permitir a manipulação das variáveis de entrada, a fim de minimizar os resultados de soda cáustica na lama vermelha e ao mesmo tempo, possa minimizar os custos envolvidos no processo. Além disso, sugere-se a utilização de técnicas adicionais, tanto para a classificação quanto para a previsão como, por exemplo, o *Support Vector Regression* (SVR), para verificar a possibilidade de atingir melhores resultados de previsão em comparados com a RNA utilizada no estudo.

REFERÊNCIAS

- ALCAN. Curso Processo *Bayer*. **Química do processo Bayer**, Canadá, 1993.
- ALCAN. Introdução ao Processo Bayer: *manual de operação*, Ouro Preto, Minas Gerais. 1994.
- ALES, V.T. **O algoritmo *sequential minimal optimisation* para resolução do Problema de support vector machine: Uma técnica para reconhecimento de padrões**. Dissertação (Mestrado em Métodos Numéricos para Engenharia) – Universidade Federal do Paraná, PR, 2008.
- ASSOCIAÇÃO BRASILEIRA DO ALUMÍNIO. Disponível em: <<http://www.abal.org.br>>. Acesso em: fev./ 2013.
- BALLINI, R.; SOARES, S; ANDRADE, G.M. **Previsão de vazões médias mensais usando redes neurais Nebulosas**. *Revista Controle & Automação* /Vol.14 no.3/Julho, Agosto e Setembro 2003.
- BAPTISTELLA, M; CUNICO, L. H. B; STEINER, M.T.A. **O Uso de Redes Neurais na Engenharia de Avaliações: Determinação dos Valores Venais de Imóveis Urbanos**. *Revista Ciências Exatas e Naturais*, Vol.9 nº 2, Jul/Dez 2007.
- BARREIROS, J.A.L; WALTER, B. JR; ANDRÉ, M.T.F; CARLOS, T.C. JR. **Estabilizador neural não-linear para sistemas de potência treinado por rede de controladores lineares**. *Revista Controle & Automação*, Vol.17 no.2/Abril, Maio e Junho 2006.
- BASSANEZI, R.C. Ensino – **Aprendizagem com modelagem matemática: uma nova estratégia**. São Paulo: Contexto, 2002.
- BELTRAMI, M. **Precificação de opções sobre ações por modelos de support vector regression**. Dissertação (Mestrado em Métodos Numéricos para Engenharia) – Universidade Federal do Paraná, PR, 2009.
- BELTRAMI, M; LOCH, G. V; SILVA, A.C.L. **Comparação das técnicas de support vector regression e redes neurais na precificação de opções**. XLII Simpósio Brasileiro de Pesquisa Operacional, 03/Ago a 03/Set. 2010.
- BLISS, N.W. **Curso Processo Bayer: Bauxita**, Canadá: *Alcan Internacional*: v, 1, 1993.
- BONCHI, F.; FERRARI, E.; JIANG, W.; MALIN, B. **Privacy, Security, and Trust in KDD**. Springer-Verlag Berlin Heidelberg, 2009.
- BRACHMAN, R. J.; ANAND, T.. **The Process of Knowledge Discovery in Databases: A First Sketch**. KDD Workshop 1994.
- BRAGA, A, P; LUDEMIR, T.B; CARVALHO, A.C.P.L. **Redes Neurais Artificiais: Teoria e Aplicações**. LTC – Livros Técnicos e Científicos / Editora S.A. – Rio de Janeiro, 2011.
- CAMPOS, M.M; SAITO, K. **Sistemas inteligentes em controle e automação de processos**. *Rio de Janeiro*: Editora Ciência Moderna Ltda, 2004.
- CAMPOS, T. E. **Técnicas de Seleção de Características com Aplicações em Reconhecimento de Faces**. Dissertação (Mestrado em Ciência da Computação – Instituto de matemática e Estatística) – Universidade de São Paulo, SP, 2001.
- CARVALHO, B. P. R. **Novas Estratégias para Detecção Automática de Vetores de Suporte em *Least Squares Support Vector Machines***. Dissertação (Mestrado em Engenharia Elétrica) – Universidade Federal de Minas Gerais, MG, 2005.
- CARVALHO, B. P. R.; BRAGA, A. P. **Estratégias neurais para treinamento de *Least Squares Support Vector Machines***. In: *Brazilian Symposium on Artificial Neural Networks*, VIII., 2004, São Luiz, Maranhão.
- CASTRO, A. A. M.; PRADO, P. P. L. **Algoritmos para reconhecimento de padrões**. *Rev. Ciênc. Exatas*, Taubaté, v. 5-8, p. 129-145, 2002.
- COELHO, R.J; SILVA, O.J; ALVES, M.T; ANDRADE, L.A; ASSIS, P.S. **Modelos de previsão da qualidade metalúrgica do coque a partir da qualidade dos carvões individuais e do coque obtido no forno-piloto de coqueificação**. *Revista. Esc. Minas*, Ouro Preto, 57(1): 27-32, jan. mar. 2005.

- CONSTANT, S; LONA, L.M.F; CALADO, V.M.A. **Predição do Comportamento Térmico de Tubos Compósitos Através de Redes Neurais.** *Ciência e Tecnologia*, vol. 14, n° 5, p. 295-300, 2004.
- CONSTANTINO, V.R.L; ARAKI, K; SILVA, D. O; OLIVEIRA, W. **Preparação de Compostos de Alumínio a partir da Bauxita: Considerações sobre alguns aspectos envolvidos em um experimento didático.** São Paulo: Universidade de São Paulo Departamento de Química Fundamental, Instituto de Química, 2002.
- CORTES, C. VAPNIK, V. **Support-Vector Networks.** *Machine Learning*, Volume 20, Número 3, Setembro/1995, pg 273-297, Editora Springer, 1995.
- CÔRTEZ, S. da C.; PROCARO, R. M.; LIFSCHITZ, S.. **Mineração de Dados – Funcionalidades, Técnicas e Abordagens. Série Monografias em Ciências da Computação.** PUC, Rio de Janeiro, 2002.
- DING, C. H.; DUBCHAK, I. **Multi-class protein fold recognition using support vector machines and neural networks.** *Bioinformatics*, v. 17, n. 4, 2001, p. 349 - 358,
- DOMINGUETE, D.H; ALMEIDA, L.P; MOREIRA, A.P; GONÇALVES, P.M; CARVALHO, S.V. **Avaliação de técnicas de regressão e de uma nova Abordagem para a otimização do processo de Desfosforação de aço no convertedor Id da v&m do Brasil.** *Tecnologia em Metalurgia e Materiais*, São Paulo, v.3. n.1, p. 31-35, jul.-set. 2006.
- ESCOBAR, E. M. **Determinação Simultânea, em linha, da Concentração de Soda Cáustica e de Aluminato em Solução de Aluminato de Sódio do Processo Bayer de Produção de Oxido de Alumínio,** Campinas, SP: Universidade Estadual de Campinas, 2004.
- FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. **From Data Mining to Knowledge Discovery in Databases.** *The American Association for Artificial Intelligence Magazine*, pag. 37-54. 1996.
- FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P.; UTHURUSAMY, R. **Advances in Knowledge Discovery & Data Mining.** 1ª ed. American Association for Artificial Intelligence, Menlo Park, Califórnia, 1996.
- FELDENS, M. A.; CASTILHO, J. M. V. **Data Mining with the Combinatorial Rule Model: an application in a health-care relational database.** In: *XXIII Conferência Latino Americana de Informática - CLEI.* Valparaiso, Chile: Universidad Technical Federico Santa Maria, 1997.
- FONTES, C.H; MENDES, M.J. **Nonlinear predictive control of an industrial slurry Reactor.** *Revista Controle & Automação*, Vol.19 no.4/Outubro, Novembro e Dezembro 2008.
- FRAWLEY, W. J.; PIATETSKY-SHAPIRO, G.; MATHEUS, C. J. **Knowledge Discovery in Databases - An Overview.** In: *Knowledge Discovery in Databases* 1991.
- GÓES, A.R.T. **Uma metodologia para a criação de etiqueta de Qualidade no contexto de descoberta de Conhecimento em bases de dados: aplicação nas áreas Elétrica e educacional.** 146 f. Tese (Doutorado em Métodos Numéricos para Engenharia) – Universidade Federal do Paraná, PR, 2012.
- GOMES, L.F.C; MONTENEGRO, S.M.G.L; VALENÇA, M.J.S. **Modelo Baseado na Técnica de Redes Neurais para Previsão de Vazões na Bacia do Rio São Francisco.** RBRH — *Revista Brasileira de Recursos Hídricos*, Volume 15 n.1 Jan/Mar 2010, 05-15.
- GRANATYR. J. **Descoberta de regras de classificação utilizando análises formal de conceitos.** PUCPR, 2011, 61 P. Dissertação de Mestrado _ Programa de Pós Graduação em Informática, Pontifícia Universidade Católica do Paraná.
- GUARNIERI, R.A. **Emprego de redes neurais artificiais e regressão Linear múltipla no refinamento das previsões de Radiação solar do modelo ETA,** São José dos Campos: INPE, 2006, 171 p. *Dissertação de Mestrado – Programa de Pós-Graduação em Meteorologia*, Instituto de Pesquisas Espaciais.
- HAN, J.; KAMBER, M. **Data Mining: Concepts and Techniques.** 2ª ed. Morgan Kauffmann Publishers, 2006.

- HAYKIN, S. **Redes Neurais: Princípios e Práticas**; trad. Paulo Martins Engel. – 2. ed. – Porto Alegre: Bookman, 2005.
- HONG, W.C; DONG, Y; ZHENG, F; LAI, C.Y. Forecasting **urban traffic flow by SVR with continuous ACO**, *Applied Mathematical Modelling* 35 (2011) 1282–1291.
- HUANG, S.C; CHUANG, P.J; WU, C.F; LAI, H.J. **Chaos-based support vector regressions for exchange rate forecasting**. *Expert Systems with Applications* 37 (2010) 8590.
- HYDRO DO BRASIL. Disponível em: <<http://www.hydro.com.br>>. Acesso em: Jun./ 2013.
- INTERNATIONAL ALUMINIUM INSTITUTE (IAI) Disponível em <<http://www.world-aluminium.org/>> Acesso em Fev./13.
- KOHAVI, R. **A study of cross-validation and bootstrap for accuracy estimation and model selection**. In: International joint Conference on artificial intelligence. [S.l.: s.n.], 1995. v. 14, p. 1137–1145.
- KRAJEWSKI, L.J; RITZMAN, L.P; MALHOTRA, M.K. **Administração de Produção e Operações** – 8ª edição – São Paulo: Pearson, 2010.
- LIMA, C. A. de M. **Comitê de máquinas: uma abordagem unificada empregando máquinas de vetores-suporte**. 378 f. Tese (Doutorado em Engenharia Elétrica) – Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, 2004.
- LIMA, E.A; FIRMINO, J.L.N; FILHO, M.F.G. **A relação da previsão da precipitação pluviométrica e casos de dengue nos estados de Alagoas e Paraíba (Nordeste do Brasil)**. *Revista Brasileira de Meteorologia*, v.23, n.3, 264-269, 2008.
- LIMA, M.M.C. **Estimativa de concentração de Material particulado em suspensão na atmosfera por meio da modelagem de rede neural artificial**. *Dissertação apresentada ao Programa de Pós-graduação em Saneamento, Meio Ambiente e Recursos Hídricos - Escola de engenharia*, Universidade Federal de Minas Gerais, Belo Horizonte, 2006.
- LOH. S. **Abordagem baseada em conceitos para descoberta do conhecimento em textos**. Tese (Doutorado em Ciência da Computação – Instituto de Informática) – Universidade Federal do Rio Grande do Sul, RS, 2001.
- LORENZI, A. **Desenvolvimento de Redes Neurais Artificiais para avaliação de estruturas de concreto através de ensaios não destrutivos**. 215 p. Tese (Doutorado em Engenharia Civil). Programa de Pós-Graduação em Engenharia Civil, UFRGS, Porto Alegre, RS. 2009.
- LUSTOSA, L; MESQUITA, M.A; QUELHAS, O; OLIVEIRA. R. **Planejamento e Controle da Produção**. Rio de Janeiro: Elsevier, 2008.
- LUZ, A.B. **Rochas e Minerais Industriais**, Editora CETEM/MCT, Rio de Janeiro 2003, 285p
- LUZ, A.B; LINS, F.A.F. **Rochas & Minerais Industriais: usos e especificações**. 2. ed. Rio de Janeiro: CETEM/MCT, 2005.
- MABILANA, H.A; FONTANA, D.C; FONSECA, E.L. **Desenvolvimento de modelo agrometeorológico espectral para estimativa de rendimento do milho na Província de Manica-Moçambique**. *Revista Ceres, Viçosa*, v. 59, n.3, p. 337-349, mai/jun, 2012.
- MACÊDO, A.N; COSTA, D.H.P; TRINDADE, S.R.S; SOUZA, J.A.S; CCARNEIRO, R.J.F.M. **Comportamento de blocos cerâmicos estruturais produzidos a partir da mistura de lama Vermelha e argila**. *Ambiente Construído*, Porto Alegre, v. 11, n. 4, p. 25-36, out./dez. 2011.
- MAHAN, B. M; MYERS, R. J. **Química: Um curso univesitário (4ª reimpressão)**, São Paulo, 2000.
- MARTINS, P.G.; LAUGENI, P.F. **Administração da Produção** – 2ª edição – São Paulo: Saraiva, 2005.
- MERCURY, J.M.R; MORAES, E.C; MONTEIRO, C.M.O. **Caracterização de Corpos cerâmicos de uma mistura de argila, Tijolos Refratários e Lama Vermelha. Parte I, estudo preliminar**. 45º Congresso Brasileiro de Cerâmica, Florianópolis SC, 2001.
- MOTA, J.F. **Um estudo de caso para a determinação do preço de venda de Imóveis urbanos via redes neurais artificiais e métodos Estatísticos multivariados**.

- Curitiba/Paraná, 2007, 86 p. *Dissertação de Mestrado – Programa de Pós Graduação em Métodos Numéricos em Engenharia*, Setor de Ciências Exatas e de Tecnologia, Universidade Federal do Paraná.
- OLIVEIRA, A.C.S; LACERDA, W.S; GONÇALVES, L.R. **Aplicação de Redes Neurais Artificiais previsão da produção de álcool.** *Ciênc. agrotec., Lavras*, v. 34, n. 2, p. 279-284, mar./abr., 2010.
- PAHL, G; BEITZ, W; FELDHUSEN J; GROTE, K. **Projeto na engenharia: Fundamentos do desenvolvimento eficaz de produtos métodos e aplicações.** São Paulo: Editora Edgard Blucher, 2005.
- PALMA, A.A; PORTUGAL, M.S. **Análise empírica da formação de Expectativas de inflação no Brasil: Uma aplicação de redes neurais artificiais a dados em painel.** *Revista Econ. contemp.*, Rio de Janeiro, 13(3): 391-437, set./dez. 2009.
- REZENDE, S.O. **Sistemas inteligentes: fundamentos e aplicações / organização. Barueri** – SP: Manole, 2005.
- ROSSI, C.R.C; RAMOS, M.A; NEGRÃO, A.M; OLIVEIRA, D.R.C. **Estudo Experimental das Propriedades Físicas e Mecânicas de Agregados para a Construção Civil Produzidos a partir de Lama Vermelha.** *50º congresso brasileiro do concreto – CBC - 2008 – 50CBC0351.*
- SANTOS, A.M; SEIXAS, J.M; PEREIRA, B.B; MEDRONHO, R.A. **Uso de redes neurais artificiais e regressão Logística na Predição de Hepatite A.** *Revista Brasil Epidemiol*, v,8, n,2, p. 117-126, 2005.
- SANTOS, R.C; SENGER, E.C. **Proposta de um algoritmo genérico baseado em R.N.A para a Proteção de distância de linhas de transmissão.** *Revista Controle & Automação/ Vol.17 no.1/Janeiro, Fevereiro e Março 2006.*
- SANTOS, V. F.; POLIDORO L.; SILVEIRA, O. F. M.; FIGUEIREDO, A. G. **Aplicação de dados multisensor (sar e etm+) no reconhecimento de padrões de uso e ocupação do solo em costas tropicais – costa Amazônica, Amapá, Brasil.** *Revista Brasileira de Geofísica*, Vol. 27(Supl. 1), 2009.
- SARTORI, L.R; IMAI, N.N; MURA, J.C; TACHIBANA, V.M. **Atributos polarimétricos de imagem radar na Inferência de parâmetros morfológicos de Macrófitas.** *Bol. Ciênc. Geod., sec. Artigos, Curitiba*, v. 18, no 1, p.138-153, jan-mar, 2012.
- SCARPARI, M.S; BEAUCLAIR, E.G.F. **Physiological model to estimate The maturity of sugarcane.** *Sci. Agric. (Piracicaba, Braz.)*, v.66, n.5, p.622-628, September/October 2009.
- SILVA, E.L; MENEZES, E.M. **Metodologia da Pesquisa e Elaboração de Dissertação. 4ª ed. rev. atual.** Florianópolis: UFSC, 2005.
- SLACK, N; LEWIS, M. **Estratégia de Operações, 2ªed.** Porto Alegre: Bookman, 2009.
- SOUZA, J.A.S. **Estudo e avaliação do uso de resíduos do processo Bayer como matéria-prima na produção de Agregados sintéticos para a construção civil.** *Tese de Doutorado apresentada ao Programa de Pós-Graduação em Engenharia de Recursos Naturais da Amazônia, ITEC, Universidade Federal do Pará – 2010.*
- STEINER, M.T.A; NETO, A.C; BRAULIO, S. N; ALVES, A.C. **Métodos estatísticos multivariados aplicados à engenharia de avaliações.** *Gest. Prod.*, São Carlos, v.15, n.1, p.23-32, jan.-abr. 2008.
- STEINER, M.T.A; NIEVOLA, J.C; SOMA, N. Y. S.T; NETO, P.J.S. **Extração de regras de classificação a partir de Redes neurais para auxílio à tomada de decisão na Concessão de crédito bancário.** *Pesquisa Operacional*, v.27, n.3, p.407-426, Setembro a Dezembro de 2007.
- SUNG, A. H.; MUKKAMALA, S. **Identifying important features for intrusion detection using support vector machines and neural networks.** *In: Symposium.(2003).*
- TAN, P. N.; STEINBACH, M.; KUMAR, V.. **Introduction to Data Mining.** Inc. Boston, MA, USA: Addison-Wesley Longman Publishing Co., 2005.
- THEODORIDIS, S., KOUTROUMBAS, K. **Pattern Recognition**, 3rd ed., Academic Press, feb. (4th 2008).

- THURAISINGHAM, B.; KHAN, L.; CLIFTON, C.; MAURER, J.; CERUTI, M. **Dependable Real-time Data Mining**. 8o IEEE International Symposium on Object- Oriented Real-Time Distributed Computing (ISORC'05), 2005.
- TOURIAN. O. K.; PADILHA, F. R. R. **Reconhecimento de variedades de soja por meio do processamento de imagens digitais usando redes neurais artificiais**. Eng. Agríc., Jaboticabal, v.28, n.4, p.759-769, out./dez. 2008.
- VAPNIK, V. **Statistical Learning Theory**. John Wiley and Sons, Inc., New York, 1995.
- VAPNIK, V. **The nature of statistical learning theory**. Springer-Verlag, New ourk, 1998.
- VAPNIK, V. N., **The Nature of Statistical Learning Theory**. New York: Springer, 1999.
- VERONEZI. C. C. D.; SIMÕES. P. W. T. A.; SANTOS. R. L.; ROCHA. E. L.; MELÃO. S; MATTOS. M.C.; CECHINEL. C. **Análise Computacional para Auxílio ao Diagnóstico de Osteoartrite de Coluna Lombar Baseado em Redes Neurais Artificiais**. Rev Bras Ortop. 2011;46(2):195-9.
- WANG, J; LI, L; NIU, D; TAN, Z. **An annual load forecasting model based on support vector regression with differential evolution algorithm**. *Applied Energy* 94 (2012) 65–70
- WEISS, S.; INDURKHYA, N. **Predictive Data Mining: a practical guide**. Morgan Kauffmann Publishers, Inc. 1998.
- WITTEN, I. H., FRANK, E. **Data Mining: Practical Machine Learning Tools and Techniques**. 2ª ed. Morgan Kauffmann Publishers, 2005.
- WU, S; AKBAROV, A. **Support vector regression for warranty claim forecasting**. *European Journal of Operational Research* 213 (2011) 196–204.
- ZAKI, M.; KARYPIS, G.; YANG, J.. **Data Mining in Bioinformatics (BIOKDD). Algorithms for Molecular Biology**, Vol. 2, 2007.

APÊNDICES

APÊNDICE A

Taxas de acertos das RNA's como Reconhecimento de Padrão

Taxa Apre	Grupos	1	3	5	7	9	11	12	13	14	16	18	20	22	23	24	25
0,1	A+BCDE	0,45	0,48	0,58	0,52	0,5	0,61	0,64	0,62	0,69	0,69	0,69	0,69	0,74	0,75	0,71	0,7
	B+CDE	0,44	0,44	0,58	0,53	0,53	0,6	0,66	0,6	0,66	0,68	0,7	0,7	0,77	0,76	0,74	0,72
	C+DE	0,47	0,49	0,57	0,51	0,6	0,65	0,68	0,64	0,69	0,69	0,67	0,68	0,72	0,76	0,73	0,72
	D+E	0,48	0,5	0,58	0,59	0,5	0,67	0,65	0,68	0,6	0,69	0,68	0,71	0,75	0,75	0,76	0,78
	Média	0,46	0,48	0,58	0,54	0,53	0,63	0,66	0,64	0,66	0,69	0,69	0,70	0,75	0,76	0,74	0,73
0,2	A+BCDE	0,46	0,5	0,59	0,63	0,68	0,76	0,79	0,77	0,75	0,67	0,7	0,67	0,72	0,72	0,69	0,74
	B+CDE	0,49	0,49	0,59	0,59	0,65	0,79	0,78	0,79	0,74	0,69	0,71	0,69	0,73	0,73	0,7	0,67
	C+DE	0,4	0,51	0,59	0,65	0,68	0,77	0,79	0,78	0,75	0,69	0,69	0,69	0,75	0,73	0,69	0,7
	D+E	0,49	0,5	0,55	0,69	0,69	0,75	0,79	0,79	0,76	0,69	0,7	0,73	0,74	0,72	0,65	0,7
	Média	0,46	0,50	0,58	0,64	0,68	0,77	0,79	0,78	0,75	0,69	0,70	0,70	0,74	0,73	0,68	0,70
0,3	A+BCDE	0,53	0,56	0,62	0,66	0,7	0,77	0,77	0,79	0,8	0,74	0,78	0,69	0,75	0,77	0,67	0,73
	B+CDE	0,55	0,59	0,62	0,65	0,65	0,78	0,78	0,8	0,79	0,69	0,79	0,74	0,75	0,77	0,63	0,7
	C+DE	0,5	0,59	0,6	0,68	0,72	0,75	0,78	0,79	0,78	0,65	0,73	0,78	0,7	0,76	0,7	0,73
	D+E	0,49	0,6	0,64	0,68	0,73	0,74	0,75	0,8	0,79	0,71	0,76	0,77	0,77	0,78	0,7	0,73
	Média	0,52	0,59	0,62	0,67	0,70	0,76	0,77	0,80	0,79	0,70	0,77	0,75	0,74	0,77	0,68	0,72
0,4	A+BCDE	0,48	0,72	0,68	0,88	0,74	0,92	0,89	0,94	0,83	0,85	0,91	0,86	0,84	0,83	0,76	0,81
	B+CDE	0,49	0,75	0,75	0,82	0,83	0,91	0,82	0,92	0,73	0,84	0,78	0,87	0,87	0,69	0,8	0,87
	C+DE	0,59	0,77	0,88	0,79	0,92	0,94	0,89	0,93	0,73	0,83	0,88	0,78	0,68	0,68	0,76	0,69
	D+E	0,56	0,78	0,74	0,75	0,85	0,88	0,92	0,92	0,8	0,73	0,76	0,87	0,81	0,79	0,78	0,77
	Média	0,53	0,76	0,76	0,81	0,84	0,91	0,88	0,93	0,77	0,81	0,83	0,85	0,80	0,75	0,78	0,79
0,5	A+BCDE	0,76	0,77	0,88	0,59	0,81	0,79	0,71	0,68	0,69	0,85	0,85	0,77	0,91	0,87	0,73	0,72
	B+CDE	0,69	0,69	0,88	0,59	0,85	0,78	0,74	0,87	0,93	0,78	0,79	0,75	0,73	0,77	0,91	0,73
	C+DE	0,82	0,74	0,83	0,68	0,89	0,82	0,76	0,68	0,77	0,77	0,67	0,91	0,64	0,89	0,81	0,79
	D+E	0,68	0,87	0,79	0,79	0,95	0,82	0,68	0,69	0,82	0,89	0,87	0,79	0,87	0,81	0,89	0,88
	Média	0,74	0,77	0,85	0,66	0,88	0,80	0,72	0,73	0,80	0,82	0,80	0,81	0,79	0,84	0,84	0,78
0,6	A+BCDE	0,59	0,74	0,9	0,87	0,87	0,76	0,85	0,79	0,78	0,93	0,92	0,79	0,57	0,84	0,92	0,81
	B+CDE	0,68	0,75	0,89	0,97	0,75	0,83	0,79	0,77	0,78	0,96	0,92	0,75	0,83	0,86	0,85	0,89
	C+DE	0,81	0,76	0,78	0,79	0,77	0,79	0,84	0,79	0,89	0,87	0,88	0,81	0,81	0,79	0,85	0,77
	D+E	0,72	0,78	0,86	0,83	0,75	0,89	0,82	0,74	0,79	0,86	0,85	0,79	0,64	0,88	0,93	0,67
	Média	0,7	0,76	0,86	0,87	0,79	0,82	0,83	0,77	0,81	0,91	0,89	0,79	0,71	0,84	0,89	0,79
0,7	A+BCDE	0,66	0,82	0,69	0,63	0,73	0,87	0,81	0,7	0,74	0,78	0,79	0,93	0,79	0,86	0,91	0,73
	B+CDE	0,68	0,8	0,68	0,69	0,84	0,86	0,57	0,73	0,79	0,91	0,86	0,83	0,83	0,87	0,88	0,65
	C+DE	0,61	0,74	0,64	0,71	0,87	0,87	0,68	0,74	0,75	0,86	0,77	0,89	0,84	0,92	0,88	0,75
	D+E	0,59	0,74	0,71	0,73	0,67	0,89	0,54	0,73	0,75	0,83	0,86	0,88	0,88	0,92	0,89	0,87
	Média	0,64	0,78	0,68	0,69	0,78	0,87	0,65	0,73	0,76	0,85	0,82	0,88	0,84	0,89	0,89	0,75
0,8	A+BCDE	0,59	0,72	0,87	0,76	0,82	0,82	0,85	0,65	0,86	0,83	0,87	0,89	0,69	0,73	0,69	0,77
	B+CDE	0,59	0,59	0,87	0,96	0,69	0,79	0,8	0,88	0,7	0,86	0,69	0,93	0,8	0,9	0,72	0,78
	C+DE	0,62	0,68	0,79	0,7	0,7	0,8	0,75	0,73	0,8	0,79	0,79	0,91	0,79	0,86	0,83	0,7
	D+E	0,5	0,91	0,74	0,79	0,64	0,68	0,94	0,59	0,83	0,79	0,93	0,89	0,67	0,87	0,79	0,77
	Média	0,58	0,73	0,82	0,80	0,71	0,77	0,84	0,71	0,80	0,82	0,82	0,91	0,74	0,84	0,76	0,76
0,9	A+BCDE	0,8	0,72	0,97	0,79	0,83	0,62	0,59	0,93	0,87	0,79	0,7	0,75	0,67	0,85	0,79	0,69
	B+CDE	0,84	0,7	0,74	0,87	0,72	0,55	0,8	0,69	0,81	0,79	0,59	0,87	0,68	0,79	0,63	0,5
	C+DE	0,79	0,87	0,74	0,69	0,89	0,75	0,93	0,95	0,84	0,63	0,88	0,71	0,93	0,59	0,7	0,55
	D+E	0,87	0,66	0,8	0,57	0,75	0,66	0,65	0,79	0,76	0,59	0,77	0,76	0,57	0,57	0,52	0,68
	Média	0,83	0,74	0,81	0,73	0,80	0,65	0,74	0,84	0,82	0,70	0,74	0,77	0,71	0,70	0,66	0,61

APÊNDICE B

MSE's das RNA's como Classificação

Neurônios	Taxa de Aprendizagem								
	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
1	0,2918	0,2929	0,2333	0,2230	0,0721	0,0962	0,1345	0,1826	0,0316
2	0,2837	0,2678	0,2311	0,0742	0,0382	0,1230	0,2183	0,0849	0,0754
3	0,2735	0,2501	0,1724	0,0605	0,0583	0,0590	0,0519	0,0892	0,0752
4	0,2222	0,1813	0,1640	0,0656	0,0192	0,0656	0,0939	0,0497	0,0444
5	0,1785	0,1767	0,1446	0,0617	0,0254	0,0225	0,1030	0,0363	0,0440
6	0,1903	0,1209	0,1141	0,0529	0,0398	0,0319	0,1743	0,0365	0,0363
7	0,2148	0,1309	0,1107	0,0383	0,1206	0,0227	0,0975	0,0483	0,0855
8	0,1681	0,1217	0,0930	0,0492	0,0640	0,0897	0,1169	0,0288	0,0712
9	0,2202	0,1058	0,0909	0,0313	0,0183	0,0487	0,0560	0,0870	0,0454
10	0,1791	0,0992	0,0756	0,0227	0,0345	0,0405	0,0127	0,0622	0,0348
11	0,1358	0,0542	0,0578	0,0081	0,0393	0,0356	0,0163	0,0547	0,1312
12	0,1175	0,0452	0,0530	0,0157	0,0779	0,0311	0,1337	0,0321	0,0838
13	0,1341	0,0473	0,0420	0,0053	0,0794	0,0521	0,0758	0,0944	0,0369
14	0,1169	0,0625	0,0441	0,0536	0,0465	0,0382	0,0591	0,0446	0,0340
15	0,1059	0,0759	0,0462	0,0430	0,0252	0,0519	0,0544	0,0324	0,0848
16	0,0976	0,0993	0,0925	0,0374	0,0339	0,0107	0,0262	0,0341	0,0983
17	0,1057	0,1057	0,0716	0,0263	0,0374	0,0107	0,0213	0,0256	0,0944
18	0,0993	0,0900	0,0557	0,0321	0,0481	0,0124	0,0340	0,0405	0,0813
19	0,0977	0,1011	0,0857	0,0336	0,0341	0,0204	0,0079	0,0136	0,0573
20	0,0931	0,0935	0,0662	0,0254	0,0419	0,046	0,0150	0,009	0,0552
21	0,0870	0,0771	0,0719	0,0467	0,1160	0,0894	0,0692	0,0822	0,1146
22	0,0653	0,0703	0,0669	0,0452	0,0568	0,0948	0,0282	0,0722	0,1002
23	0,0601	0,0756	0,0529	0,0678	0,0295	0,0259	0,0123	0,0298	0,1049
24	0,0705	0,1011	0,1064	0,0509	0,0323	0,0140	0,0122	0,0618	0,1253
25	0,0738	0,0891	0,0771	0,0505	0,0524	0,0525	0,0687	0,0610	0,1627

APÊNDICE C

MSE's das RNA's como Previsão

Neurônios	Taxa de Aprendizagem								
	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
1	0,03182	0,03200	0,03244	0,02839	0,033519	0,031918	0,03191	0,03191	0,028438
2	0,03240	0,03049	0,03105	0,03378	0,033975	0,029295	0,03209	0,03111	0,030994
3	0,03235	0,03192	0,03038	0,03143	0,030732	0,030491	0,03345	0,03463	0,032561
4	0,03402	0,03138	0,03243	0,02979	0,032472	0,031247	0,02870	0,03278	0,033095
5	0,03254	0,03212	0,03164	0,03208	0,030823	0,029473	0,02675	0,02963	0,029691
6	0,03119	0,03130	0,03068	0,03346	0,030566	0,030064	0,02528	0,02949	0,029376
7	0,03236	0,03162	0,03394	0,02730	0,026553	0,030479	0,03000	0,03093	0,032371
8	0,03098	0,03156	0,03175	0,03666	0,029374	0,025947	0,02640	0,03029	0,030697
9	0,03361	0,03063	0,03172	0,02529	0,023548	0,028411	0,02697	0,02716	0,029943
10	0,03319	0,03220	0,03184	0,02327	0,028999	0,031377	0,02351	0,02999	0,042479
11	0,03283	0,03254	0,03215	0,02828	0,024735	0,025401	0,02374	0,02916	0,024839
12	0,03161	0,03178	0,03001	0,02580	0,030074	0,020248	0,01811	0,02215	0,032741
13	0,03098	0,03153	0,03459	0,03014	0,027484	0,024911	0,02329	0,02985	0,023261
14	0,03258	0,0315	0,03175	0,02962	0,026005	0,021849	0,03113	0,0221	0,027439
15	0,03139	0,03286	0,03183	0,02195	0,016797	0,017319	0,02088	0,02146	0,022987
16	0,03084	0,03030	0,03215	0,02566	0,01854	0,018142	0,02617	0,02372	0,023004
17	0,03234	0,03164	0,03467	0,02233	0,02106	0,021823	0,01863	0,02455	0,023517
18	0,03155	0,03315	0,03179	0,02343	0,02317	0,023047	0,02499	0,02196	0,023514
19	0,03317	0,02999	0,01786	0,01986	0,014724	0,025585	0,01961	0,01569	0,020761
20	0,03337	0,03050	0,02867	0,02752	0,026117	0,025585	0,01831	0,01877	0,01616
21	0,03142	0,02976	0,01885	0,01819	0,016822	0,01693	0,01706	0,01652	0,021357
22	0,03299	0,03360	0,01605	0,02622	0,020828	0,017682	0,02199	0,01795	0,020702
23	0,03297	0,03227	0,01815	0,02458	0,002157	0,021806	0,02203	0,02216	0,015398
24	0,03035	0,02952	0,02602	0,01511	0,017807	0,016229	0,00166	0,01651	0,025222
25	0,03253	0,03121	0,01771	0,01814	0,015484	0,016676	0,01011	0,01835	0,02348