



**UNIVERSIDADE DO VALE DO RIO DOS SINOS
CIÊNCIAS EXATAS E TECNOLÓGICAS
PROGRAMA INTERDISCIPLINAR DE PÓS-GRADUAÇÃO
EM COMPUTAÇÃO APLICADA**

Rafael Schardosin Silva

**PREVISÃO DE VENDA DE PRODUTOS EM UMA
INDÚSTRIA DE TELECOMUNICAÇÃO
UTILIZANDO REDES NEURAIAS ARTIFICIAIS**

SÃO LEOPOLDO,

2010



**UNIVERSIDADE DO VALE DO RIO DOS SINOS
CIÊNCIAS EXATAS E TECNOLÓGICAS
PROGRAMA INTERDISCIPLINAR DE PÓS-GRADUAÇÃO
EM COMPUTAÇÃO APLICADA**

Rafael Schardosin Silva

**PREVISÃO DE VENDA DE PRODUTOS EM UMA
INDÚSTRIA DE TELECOMUNICAÇÃO
UTILIZANDO REDES NEURAIAS ARTIFICIAIS**

Dissertação apresentada ao Programa de Pós-Graduação em Computação Aplicada da Universidade do Vale do Rio dos Sinos como requisito parcial para obtenção do título de Mestre em Computação Aplicada.

Orientador: Prof. Dr. João Valiati

SÃO LEOPOLDO,

2010

A Débora Ferraz Galo,
pelo companheirismo e apoio incondicional.

AGRADECIMENTOS

Ao meu orientador Prof. Dr. João Francisco Valiati por todos os ensinamentos, apoio e confiança na realização deste trabalho, além de ser um grande ser humano, pela amizade criada, tornando o percurso muito mais agradável.

Ao Prof. Dr. Adelmo Luis Cechin, grande mestre, cujos ensinamentos e experiências transmitidos serão levados por mim pelo resto da vida. Que descanse em paz.

A todos os professores do PIPCA, em especial ao Prof. Dr. Luis Paulo Luna de Oliveira, pelas aulas descontraídas, mas com muito conhecimento transmitido, e pela solicitude em ajudar e contribuir a qualquer momento.

Aos meus pais e irmão pela confiança e amor, sempre depositados em mim e pelo exemplo de caráter demonstrado, o que norteia minha vida.

A minha noiva, pelo amor, apoio e compreensão, mesmo nos momentos de ausência necessários para realização do trabalho.

Ao meu sogro e sogra, por serem pessoas sensacionais, os quais já considero como uma extensão da minha família.

A todos os meus amigos, por mostrarem a vida em uma perspectiva muito mais divertida e por estarem ao meu lado em todos os momentos.

À direção e aos funcionários da empresa onde o trabalho foi realizado. Agradeço por terem facilitado a realização do trabalho e pela bolsa concedida, pois sem ela a realização desse curso de mestrado teria sido muito difícil.

A todos que de alguma forma contribuíram para a realização desta pesquisa.

RESUMO

Este trabalho tem por objetivo apresentar uma proposta para a previsão de vendas de produtos, com fundamentação teórica, por meio da utilização de Redes Neurais Artificiais, utilizando como estudo de caso uma indústria que desenvolve produtos para o ramo de telecomunicação.

Atualmente, realizar previsões de venda nas empresas é fundamental para reduzir seus custos com gastos desnecessários em recursos humanos e materiais e aumentar sua liquidez, sem perder a qualidade que os clientes estão acostumados, no sentido de evitar atrasos nos prazos de entrega dos produtos.

O problema abordado nesta dissertação é importante para quase todas as áreas das indústrias, uma vez que a correta previsão de vendas permite às indústrias uma melhor organização de seu setor produtivo, permitindo a antecipação da quantidade ideal de matéria-prima a ser adquirida, o alinhamento de sua linha de produção, de modo a não ocorrerem alterações bruscas em seus *layouts* de fábrica, e um maior controle de seus níveis de estoque, reduzindo excessos sem ocasionar a falta de produtos.

O estudo realizado e a solução proposta são baseados em Redes Neurais Artificiais aplicadas à previsão de venda de produtos com base no histórico de vendas passadas e na inserção de variáveis auxiliares que contribuem na caracterização do problema. Os resultados obtidos nesse estudo se demonstram promissores uma vez que se alcançou uma queda significativa no erro de previsão obtido pelo método usualmente empregado.

Algumas das motivações para a realização do trabalho foram as limitações do método atualmente utilizado, cujo não tem a capacidade compreender padrões sazonais e não lineares, considerando para o período presente, somente a média de um período previamente definido, trazendo para a empresa uma elevada taxa de erro.

O principal objetivo da pesquisa foi a de apresentar uma alternativa aos métodos estatísticos tradicionais, os quais são atualmente utilizados pelas indústrias como ferramenta de previsão de vendas. Como consequência, críticas sobre o método atualmente utilizado são apresentadas, assim como são realizados comparativos, mostrando uma significativa melhoria nas previsões utilizando o método proposto por esse trabalho.

ABSTRACT

This work has the objective to present a methodology to predict product sales using Artificial Neural Networks, as a study case was treated the periodic sales volume of an industry which develops products to the telecommunication area.

Nowadays, realize sales prediction in the companies is crucial to reduce the costs with human resources, materials and increase the liquidity, without loose the quality which the clients are familiarized, avoiding delays in the products deliveries.

The problem which this dissertation tackles is important to several kinds of industries, once a precise sales prediction allow the industries to organize themselves in a better way. It allows the company to know anticipated the material to be acquired, align their productions lines, avoiding an abrupt change in the factory layouts and offers a better control of their stocks.

The study realized and the solution proposed are based on Artificial Neural Networks, applied to predict product sales using the sales history and the insertion of auxiliary variables which can contribute to characterize the problem. The results obtained in this study are promising, once it achieved an expressive reduction of the error rate when compared to the method currently applied by the industry studied.

Some of the motivators for doing the work were the limitations of the method currently used, which lacks the ability to understand seasonal and non-linear patterns, considering to the current period, just the average of a period previously defined, bringing to the company a very high error rate.

The main intention of this research was to present an alternative solution to the traditional statistical methods, which are currently used by the industries as a way to predict products sales. As a consequence, criticisms are presented to the current method used, as well a comparative results evaluation, showing a significant improvement in the predictions using the method proposed by this work.

LISTA DE FIGURAS

Figura 2.1: Componentes de um sistema de informação.....	17
Figura 2.2: Modelo Conceitual de um SAD.....	19
Figura 2.3: Sistema de Gestão Empresarial.....	20
Figura 3.1 As partes de um neurônio.	31
Figura 3.2: Relação entre a frequência de pulsos de um neurônio e a intensidade do estímulo de longa duração.	32
Figura 3.3: Neurônio Artificial ou Unidade de Processamento.	33
Figura 3.4: Exemplo de Rede <i>Perceptron</i>	35
Figura 3.5: Rede Neural MLP com uma camada escondida.	37
Figura 4.1: Diagrama da Metodologia Proposta	52
Figura 4.2: Arquitetura da RNA do Experimento 1	60
Figura 4.3: Arquitetura da RNA do Experimento 2	62
Figura 4.4: Arquitetura da RNA do Experimento 3	64
Figura 5.1: Erro da RNA por <i>fold</i> obtido no Experimento 1.....	65
Figura 5.2: Erro do método tradicional por <i>fold</i> obtido no Experimento 1.....	66
Figura 5.3: Erro da RNA por <i>fold</i> obtido no Experimento 2.....	67
Figura 5.4: Erro da RNA por <i>fold</i> obtido no Experimento 3.....	68

LISTA DE TABELAS

Tabela 2.1: Exemplo do MA de ordem 3.....	28
Tabela 3.1: Funções de Ativação mais utilizadas.....	34
Tabela 4.1: Exemplo de informações de produtos utilizadas no Previsão de Vendas.....	53
Tabela 4.2: Resultado da Previsão de Vendas utilizado atualmente na empresa	53
Tabela 4.3: Exemplo de dados utilizados pelo <i>Forecasting</i>	56
Tabela 4.4: Exemplo de dados utilizados no Experimento 1	59
Tabela 4.5: Exemplo de dados utilizados no Experimento 2	61
Tabela 4.6: Exemplo de dados utilizados no Experimento 3	63

LISTA DE ABREVIATURAS E SIGLAS

ARIMA	– Autoregressive Moving Average
BD	– Banco de Dados
CRM	– Customer Relationship Management
EIS	– Executive Information System
ERP	– Enterprise Resource Planning
IA	– Inteligência Artificial
MA	– Moving Average
MAPE	– Mean Absolute Percentage Error
MLP	– Multi-Layer-Perceptron
OO	– Orientação a Objetos
RNA	– Rede Neural Artificial
SAD	– Sistema de Apoio à Decisão
SE	– Sistemas Especialistas
SIG	– Sistema de Informação Gerencial
ST	– Sistemas Transacionais
UNISINOS	– Universidade do Vale do Rio dos Sinos
VLSI	– Very Large Scale Integration

SUMÁRIO

1. INTRODUÇÃO	12
1.1. OBJETIVO PRINCIPAL	13
1.2. OBJETIVOS ESPECÍFICOS	13
1.3. ORGANIZAÇÃO DO TRABALHO	14
2. SISTEMAS DE INFORMAÇÃO E PREVISÃO DE VENDAS.....	16
2.1. SISTEMAS DE INFORMAÇÃO.....	16
2.1.1. Sistemas de Apoio à Decisão.....	17
2.1.2. ERP	19
2.2. PREVISÃO DE VENDAS	21
2.2.1. Tipos de Previsão.....	24
2.3. MOVING AVERAGE.....	27
3. REDES NEURAIS ARTIFICIAIS.....	30
3.1. DEFINIÇÃO.....	30
3.2. O NEURÔNIO BIOLÓGICO	31
3.3. NEURÔNIO ARTIFICIAL	32
3.4. O <i>PERCEPTRON</i>	34
3.5. <i>PERCEPTRON</i> DE MÚLTIPLAS CAMADAS	36
3.5.1. Arquitetura dos <i>Perceptrons</i> de Múltiplas Camadas	36
3.6. TREINAMENTO DE REDES NEURAIS ARTIFICIAIS.....	37
3.6.1. O Algoritmo Backpropagation.....	39
3.7. PROPRIEDADES E CAPACIDADES DAS RNAS	41
3.8. TRABALHOS CORRELATOS	44
3.8.1. Previsão de Demanda de Energia Elétrica	45
3.8.2. Previsão de Vendas	47
3.8.3. Observações sobre os Trabalhos.....	50
4. MODELO ATUAL E EXPERIMENTOS PARA PREVISÃO DE VENDAS.....	51
4.1. MODELO ATUAL	52
4.2. FONTE DE DADOS	54
4.3. FERRAMENTAS	54

4.4. COLETA DOS DADOS.....	55
4.5. TÉCNICAS PARA AVALIAÇÃO DOS RESULTADOS	57
4.5.1. K-fold Cross Validation.....	57
4.5.2. MAPE	58
4.6. EXPERIMENTO 1	58
4.6.1. Modelagem e Treinamento da RNA	59
4.7. EXPERIMENTO 2	60
4.7.1. Modelagem e Treinamento da RNA	61
4.8. EXPERIMENTO 3	62
4.8.1. Modelagem e Treinamento da RNA	63
5. RESULTADOS	65
5.1. EXPERIMENTO 1	65
5.2. EXPERIMENTO 2	67
5.3. EXPERIMENTO 3	68
6. CONCLUSÕES.....	70
REFERÊNCIAS	72
ANEXO 1 – EXEMPLO DE ENTRADAS E SAÍDAS DO EXPERIMENTO 1	75
ANEXO 2 – EXEMPLO DE ENTRADAS E SAÍDAS DO EXPERIMENTO 2	76
ANEXO 3 – EXEMPLO DE ENTRADAS E SAÍDAS DO EXPERIMENTO 3	77
ANEXO 4 – ERRO DA REDE NOS EXPERIMENTOS.....	78

1.INTRODUÇÃO

Realizar previsões de demanda é importante para auxiliar na determinação dos recursos necessários para a empresa. Em tempos de abertura de mercados, essa atividade torna-se fundamental. Os mercados que podem ser acessados pela empresa, assim como a concorrência que os disputa, mudam continuamente, exigindo novas previsões de demanda em períodos mais curtos (MAKRIDAKIS, 1998).

Na indústria, conseguir prever a venda dos produtos é fator de extrema importância para que a empresa consiga organizar seus recursos humanos e materiais, de forma a atender seus clientes pontualmente, primando pela qualidade e organizando suas atividades de modo a ter o menor custo possível.

Previsões de demanda são elaboradas utilizando métodos quantitativos, qualitativos ou combinações de ambos. Métodos quantitativos baseiam-se na análise de séries temporais (dados que descrevem a variação da demanda ao longo do tempo). Métodos qualitativos baseiam-se em opiniões de especialistas, sendo vulneráveis a tendências que podem comprometer a confiabilidade de seus resultados. Métodos qualitativos costumam apresentar um baixo grau de precisão, sendo, mesmo assim, os mais utilizados por empresas (MENTZER e COX, 1984).

Devido à dependência que as empresas hodiernas têm em relação aos sistemas de informação, os dados históricos são armazenados e utilizados com maior facilidade. Isso permite a obtenção e uso de informações importantes que podem contribuir na tomada de decisão, de forma a incentivar as empresas a aumentarem a utilização do modelo de previsão quantitativo, entendendo a dinâmica de venda dos seus produtos, e assim viabilizar uma previsão com maior acerto.

Segundo Laudon e Laudon (1999), um sistema de informação pode ser definido como um conjunto de componentes inter-relacionados trabalhando juntos para coletar, recuperar, processar, armazenar e distribuir informação com a finalidade de facilitar o planejamento, o controle, a coordenação, a análise e o processo decisório em empresas e outras organizações.

Esta dissertação tem o objetivo de apresentar e propor a utilização de Redes Neurais Artificiais (RNAs) na busca da previsão de venda de produtos através das informações históricas das vendas dos produtos de uma empresa do ramo de telecomunicação.

As RNAs podem ser utilizadas em conjunto com os métodos clássicos de previsão de vendas utilizados nos sistemas de informação, possibilitando agregar no momento da previsão

mais uma informação valiosa originada pela RNA, em que o especialista poderá, além da utilização do método clássico baseado em histórico (quantitativo) e da decisão baseada em informações e experiência (qualitativo), contar também com o conhecimento adquirido pela RNA através da dinâmica de vendas passadas, de maneira a realizar previsões de venda.

Devido ao grande volume de dados armazenados pelas corporações hoje em dia, utilizar sistemas inteligentes para adquirir informações baseadas em fatos históricos é um diferencial muito valioso para as empresas, tornando-as mais competitivas e preparadas para lidar com fatores que determinarão seu sucesso. RNAs tem tido, nos últimos anos, um crescimento considerável, mas ainda pouco utilizado na indústria; e uma das intenções desta dissertação é divulgar e demonstrar a possibilidade da aplicação da Inteligência Artificial (IA), neste caso, das RNAs, em uma das áreas mais importantes da indústria: a previsão de venda de produtos em que quanto mais próximo for a previsão da realidade, melhor será a organização da empresa quanto a sua produção, incluindo compras de matéria prima que na maioria das vezes tem origem no exterior, o que exige o mínimo de erro para que os prazos de entrega sejam cumpridos.

1.1.OBJETIVO PRINCIPAL

O objetivo principal desta dissertação é modelar e avaliar a utilização de Redes Neurais Artificiais na previsão de venda de produtos, considerando como estudo de caso uma indústria de suprimentos de equipamentos para telecomunicação, buscando realizar um comparativo entre as RNAs e os métodos tradicionais e averiguar a possibilidade de redução do erro nas previsões.

1.2.OBJETIVOS ESPECÍFICOS

Como objetivos mais específicos este trabalho se propõe a:

- Realizar a revisão do estado da arte de trabalhos correlatos a este.

- Modelar um conjunto de RNAs para o problema e avaliar os melhores resultados obtidos com cada arquitetura mediante dados de teste.
- Comparar os melhores modelos utilizando RNAs com o modelo quantitativo utilizado pela empresa.
- Avaliar que outras variáveis contribuiriam para obtenção de um resultado melhor na utilização de RNAs.
- Identificar a viabilidade da utilização de RNAs como uma alternativa complementar, o que contribui como solução adicional às já existentes.

1.3. ORGANIZAÇÃO DO TRABALHO

No capítulo 2 será apresentado o referencial teórico sobre o tema abordado, serão apresentados os modelos de Previsão de Vendas utilizados hoje em dia pelas empresas, o funcionamento destes e também os Sistemas de Informação onde as Previsões de Vendas estão inseridas.

No capítulo 3 são apresentadas as Redes Neurais Artificiais, que nesta dissertação é o método utilizado na tentativa de aumentar a capacidade de previsão de venda de produtos pela empresa estudada, o seu funcionamento e todos os aspectos que compõem uma RNA, assim como trabalhos correlatos que justificam a escolha de RNA na solução de problemas dessa natureza.

Na sequência, no capítulo 4 é apresentado o modelo atual de previsão, expondo seu funcionamento, a fonte dos dados obtidos, as ferramentas utilizadas, bem como são detalhados três experimentos utilizando RNAs e a evolução deles, em que são apresentados a definição da arquitetura assim como os passos realizados no treinamento.

No capítulo 5, os resultados obtidos nos experimentos realizados são apresentados, assim como é realizado um comparativo do método utilizado atualmente pela empresa com os experimentos e também é realizado um comparativo entre os experimentos, identificando evoluções de um modelo para outro.

Por fim, no capítulo 6 são apresentadas as conclusões obtidas após a realização do trabalho, analisando a estrutura atual de previsão de vendas utilizada pela empresa, assim

como sobre uma possível estrutura utilizando RNAs para o mesmo objetivo. Também são apresentadas expectativas para trabalhos futuros que podem ser realizados nessa área.

2.SISTEMAS DE INFORMAÇÃO E PREVISÃO DE VENDAS

Nesse capítulo será apresentada a fundamentação teórica necessária para compreensão da pesquisa realizada, serão apresentados os sistemas de informação em que o tema do trabalho está inserido, assim como os tipos de previsão de vendas utilizados, qual o seu objetivo, o seu funcionamento e o método quantitativo que a empresa utiliza em seu cotidiano.

2.1.SISTEMAS DE INFORMAÇÃO

Ao longo dos anos os Sistemas de Informação (SI) se mostraram fundamentais nas organizações, cujo nível de dependência é tão elevado que não conseguem mais imaginar os seus processos sem o apoio de um Sistema de Informação, utilizando-o tanto nos processos diários como no apoio à tomada de decisão. Segundo Laudon e Laudon (1999), muitas empresas veem os sistemas de informação como cruciais para a sustentação do negócio a longo prazo. Eles definem um SI como:

"Um conjunto de componentes inter-relacionados trabalhando juntos para coletar, recuperar, processar, armazenar e distribuir informação com a finalidade de facilitar o planejamento, o controle, a coordenação, a análise e o processo decisório em empresas e outras organizações."

Um Sistema de Informação consiste em: (a) uma entrada, que são os dados brutos coletados da empresa; (b) o processamento, que é o passo em que os dados são processados e transformados em informação; e (c) a saída, que é o meio que o sistema de informação utiliza para disponibilizar essas informações para as pessoas que as necessitam. Além desses três passos, existe também a retroalimentação, que é o retorno da informação para um dos passos anteriores, buscando uma melhora ou correção da informação. Abaixo, na Figura 2, é apresentado o funcionamento de um SI com os seus componentes.

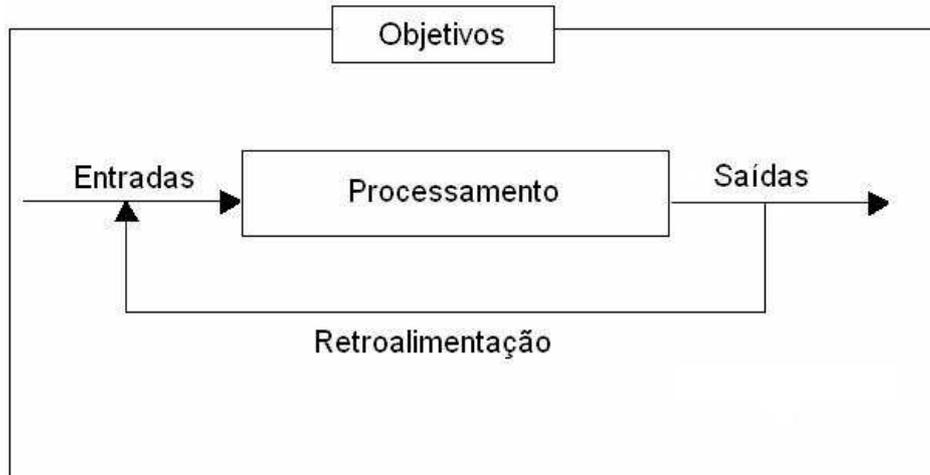


Figura 2.1: Componentes de um sistema de informação

Existem diversas classificações para os Sistemas de Informação com o objetivo de diferenciar esses sistemas para as diversas finalidades que ele é utilizado. Os principais sistemas de informação são os Sistemas de Informação Gerencial (SIG), Sistemas de Apoio a Decisão (SAD), Sistemas Especialistas (SE), *Executive Information System* (EIS), *Enterprise Resource Planning* (ERP) e o *Customer Relationship Management* (CRM). Alguns desses tipos de Sistemas de Informação serão tratados nos itens a seguir.

Cabe salientar que não existe uma fronteira definida entre os diversos tipos de sistemas de informação, como por exemplo, um EIS não deixa de ser um tipo de sistema de apoio a decisão.

É difícil de imaginar uma empresa nos dias atuais sem um sistema de informação, pois este é parte estratégica e fundamental das empresas e, com certeza, um dos pontos mais expressivos na busca pelo sucesso nos negócios. Com o volume das informações cada vez maior, ter precisão e velocidade na resposta a questões cruciais do negócio são fatores extremamente importantes para uma empresa tornar-se competitiva. A não utilização de um sistema de informação torna-se quase inadmissível.

2.1.1. Sistemas de Apoio à Decisão

No início dos anos 70, com a forte crescente dos sistemas de informação, surgiu a necessidade de um sistema que apoiasse a tomada de decisão, trazendo informações mais precisas, aumentando a qualidade das informações e a velocidade na tomada de decisão.

Segundo Turban, McLean e Wetherbe (2004), Sistema de Apoio a Decisão é um sistema de informação baseado em computador que combina modelos e dados, numa tentativa de solucionar problemas semiestruturados com grande envolvimento por parte do usuário.

Devido à complexidade na realização de negócios, fazer escolhas certas é um fator crucial para as empresas. Sem o apoio de um SAD é muito provável que os gerentes não consigam obter o máximo em resultados ao tomar uma decisão somente à luz de sua visão, ocasionando possíveis perdas para a empresa. O SAD vem ao encontro dos tomadores de decisão fornecendo informações precisas sobre o seu negócio, propondo soluções e permitindo aos tomadores de decisão a possibilidade de simular situações. O SAD geralmente utiliza como fonte de dados os Sistemas Transacionais (ST), os Sistemas de Informações Gerenciais (SIG) e fontes externas.

Segundo Stair (1998), os Sistemas de Apoio a Decisão podem:

- Manipular grandes volumes de dados.
- Obter diferentes fontes de dados (dados de computadores pessoais, *mainframes*, redes de computadores).
- Proporcionar flexibilidade de relatórios e de apresentações (formatos variados de relatórios produzidos em tela de vídeo ou de forma impressa).
- Possuir orientação tanto textual quanto gráfica (de acordo com a preferência do gerente).
- Realizar análises e comparações complexas e sofisticadas.
- Dar suporte às abordagens de otimização, satisfação e heurística (em problemas menores o SAD tem capacidade de encontrar a solução ótima; no caso de problemas complexos, são utilizadas abordagens de satisfação e heurística e assim o computador pode encontrar uma solução muito boa – não obrigatoriamente a melhor).
- Análises de simulações e por metas. A análise de simulações é o processo de fazer modificações hipotéticas nos dados do problema e observar os impactos nos resultados, enquanto a análise de alcance das metas é um processo inverso em que a partir do resultado almejado, são determinados os dados requeridos do problema.

Segundo Turban, McLean e Wetherbe (2004), todos os SADs consistem pelo menos em componentes de gestão de dados, interface de usuário e gestão de modelos e, é claro, de

usuários. Alguns poucos SADs mais avançados também contêm um componente de gestão do conhecimento. Abaixo, na Figura 2.2, pode-se visualizar o modelo conceitual de um SAD.

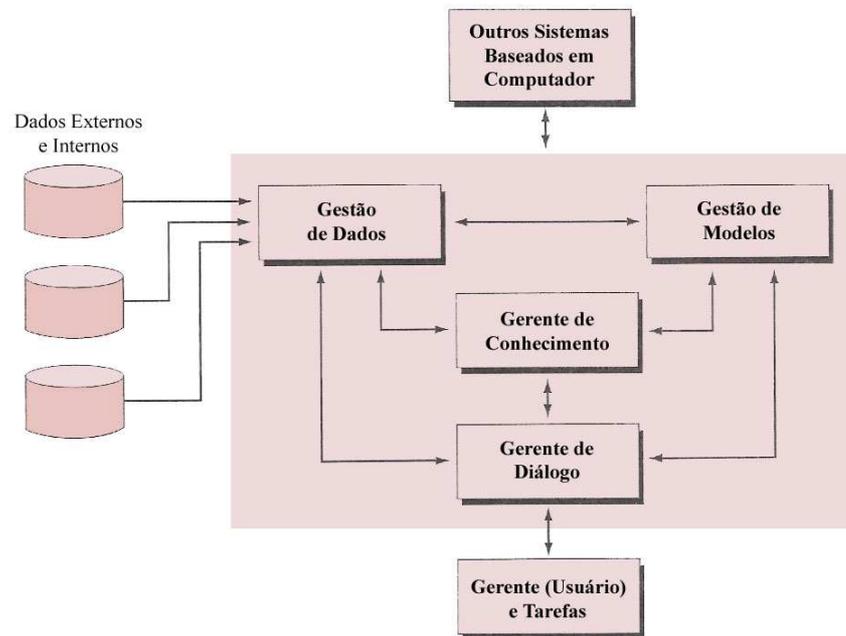


Figura 2.2: Modelo Conceitual de um SAD, adaptado de Turban, McLean e Wetherbe (2004).

A figura 2.2 apresenta a interação entre o SAD com suas fontes de dados, sendo elas tanto os dados externos e internos como também outros sistemas interligados e os usuários. No caso dos outros sistemas e os usuários, eles além de fornecerem dados, também são os clientes do SAD, podendo utilizar as informações contidas.

2.1.2.ERP

Antes dos anos 50 todas as informações dentro das organizações eram controladas no papel, isso propiciava uma grande ineficiência no controle e obtenção das informações. Ao surgirem os softwares, no final da década de 50, as empresas iniciaram processos de implantação de sistemas em todas as áreas que os necessitavam, buscando redução de custos e aumento da eficiência. Os softwares, na época, eram desenvolvidos somente com a visão funcional de cada área. Por exemplo, existiam softwares que realizavam o controle da produção e softwares que controlavam os estoques, mas mesmo essas duas áreas tendo

ligações e uma impactando diretamente na outra, as informações não eram compartilhadas entre esses dois softwares, dificultando a visão e controle global dos processos da empresa.

Ao identificar a necessidade de que os processos fossem integrados, surgiu o que foi definido pelos teóricos como ERP, que é a unificação de todos os processos da empresa em uma única arquitetura de software, possibilitando um melhor planejamento e gestão geral dos recursos e visualização dos impactos de uma área sobre a outra em tempo real.

O principal objetivo do ERP é integrar todos os departamentos e funções da empresa em um sistema unificado de informática, com capacidade de atender a todas as necessidades da organização (TURBAN, MCLEAN e WETHERBE, 2004).

Para Laudon e Laudon (1999), o sistema de gestão empresarial tem o potencial de integrar os processos-chave da organização em um sistema único, com os seus limites permeáveis em relação aos clientes e aos vendedores. A Figura 2.3 apresenta a lógica conceitual que está por trás dos sistemas ERP.



Figura 2.3: Sistema de Gestão Empresarial, adaptado de Laudon e Laudon (1999).

A implantação de um ERP tem custo e risco elevados para as organizações. Mas ao ser implantando exige que seja feito uma re-engenharia de processos na empresa, buscando que todos os funcionários trabalhem de acordo com a metodologia proposta, reduzindo drasticamente os pontos ineficientes do processo.

Uma das características dos ERPs é o seu desmembramento em componentes individuais, possibilitando assim que sejam ajustados ou que sejam acrescentadas novas funcionalidades conforme a realidade da empresa. Dessa forma, também a atualização do ERP ocorre de uma maneira muito mais simples, sem afetar toda a organização, isso significa que, caso exista uma alteração no componente financeiro, a atualização ocorra somente para esse componente, sem a necessidade de se atualizar todo o ERP.

Apesar de todas as correntes favoráveis ao ERP, ainda existem muitos casos de fracasso na sua implantação, principalmente em função da falta de um planejamento adequado e não levantamento do impacto que ele terá sobre o processo atual da empresa.

De acordo com Turban, McLean e Wetherbe (2004), para que a implantação de um ERP seja bem sucedida, no diálogo inicial devem ser incluídos os seguintes fatores:

- As expectativas dos clientes.
- Os recursos e falhas do produto ERP utilizado.
- O nível de mudança que o cliente precisará enfrentar para o ajuste do sistema.
- O nível de comprometimento, na empresa cliente, com a concretização do projeto.
- O ajuste entre a organização e cultura do cliente e a organização e cultura do projeto.
- Os riscos decorrentes das políticas vigentes no âmbito da empresa cliente.
- As habilidades, responsabilidades e funções do consultor (quando houver).

2.2.PREVISÃO DE VENDAS

A previsão de vendas é fundamental para as empresas, tanto no aspecto financeiro como na satisfação dos seus clientes. Conseguir prever a demanda dos seus produtos de forma correta permite que a empresa consiga organizar a aquisição de matéria prima, sua linha de produção, assim como todos os aspectos ligados ao produto, de forma a evitar imprevistos, gerando um custo desnecessário e até mesmo atrasos na entrega de seus produtos.

As empresas que têm em sua cultura nunca atrasar pedidos, utilizando para essa finalidade estoques de segurança, têm a vantagem de seu cliente ter a certeza que terá o produto no prazo estipulado. Por outro lado, manter estoque de segurança gera custo, tanto na compra de matéria-prima desnecessária como custo para manter o estoque na empresa. E assim, enquanto esse valor poderia estar sendo reinvestido, o dinheiro é gasto em um custo desnecessário, diminuindo a liquidez da empresa por ela não conseguir prever adequadamente a demanda de seus produtos.

A previsão de vendas é uma forma de estimar a demanda futura por meio da antecipação do que os compradores provavelmente farão em determinado conjunto de condições. A previsão de vendas fornece uma estimativa da demanda dos produtos

individuais, a qual, quando integrada ao plano agregado de produção, gera para a empresa a quantidade e os dados dos produtos finais individuais. O processo de monitorar e de integrar essas informações é denominado gestão da demanda (CAVALHEIRO, 2003).

A Previsão de vendas da empresa é o nível esperado de vendas em determinado período, baseado no plano de marketing escolhido e no ambiente de marketing assumido (KOTLER, 1998).

Segundo Corrêa, Gianesi e Caon (2001), a gestão de demanda abrange cinco áreas principais, destacando-se a previsão de demanda, a comunicação com o mercado, a influência sobre a demanda, a promessa de prazos de entrega, além da priorização e alocação de pedidos. Essas áreas podem ser especificadas da seguinte maneira:

- Habilidade para prever a demanda: é importante que a empresa consiga antecipar a demanda com alguma precisão. Isso pode envolver a manutenção de uma base de dados históricos de vendas, assim como informações das variações ocorridas ao longo do tempo. A empresa pode usar modelos matemáticos adequados que ajudam a explicar o comportamento da demanda.
- Canal de comunicação com o mercado: os vendedores ou representantes têm um papel muito importante de trazer informações dos clientes e do mercado para a empresa, de forma contínua e permanente, contribuindo para a gestão de demanda.
- Poder de influência sobre a demanda: além de prever o comportamento da demanda, é fundamental que a empresa procure influenciá-lo, por meio de propaganda ou por meio de promoção.
- Habilidade de prometer prazos: a garantia do desempenho na confiabilidade de entregas é responsabilidade de quem faz a gestão de demanda.
- Habilidade de priorização e alocação: o atendimento dos pedidos realizados pelos clientes é o principal objetivo da gestão de demanda.

Entre os sistemas de informação utilizados pelas organizações, a previsão de demanda se encontra nos Sistemas de Apoio a Decisão (SAD), sistemas que geralmente se encontram dentro dos Sistemas Integrado de Gestão (*Enterprise Relationship Management* – ERP) e que utilizam informações do ERP na busca de informações relevantes que apoiem a tomada de decisão.

Segundo Nahmias (1993), algumas características da Previsão de Demanda podem ser descritas, tais como:

- Previsões são em geral erradas: por mais estranho que possa parecer, provavelmente esta é a mais ignorada e importante propriedade de quase todos os métodos de previsão. Comumente, as previsões são feitas com bases em informações conhecidas. As necessidades de recurso para a programação de produção e compras requer modificações se a Previsão de Demanda estiver imprecisa. O sistema de planejamento deve ser suficientemente robusto para ter capacidade de antecipar erros de previsão.
- Uma boa previsão é mais que um simples número: já que as previsões são usualmente erradas, uma boa previsão também inclui alguma medida para antecipar o erro da previsão.
- Previsões agregadas são mais exatas: de acordo com a estatística, a variação da média de um conjunto de valores é menor que a variação de cada variável isolada. Esse fenômeno também se aplica na Previsão de Demanda. O erro em uma previsão de uma linha inteira de produtos é menor que o erro da uma previsão de um item individualmente.
- Quanto mais longo for o horizonte da previsão, menos exata ela será: trata-se de uma propriedade que é quase intuitiva. Pode-se prever o dia de amanhã com mais exatidão do que um dia no próximo ano.
- Previsões não deveriam ser usadas para excluir uma informação conhecida: algumas informações podem não estar presente na história e no comportamento de uma demanda, mas não devem ser excluídas. Por exemplo, a empresa pode planejar uma venda promocional especial para um item em particular que fará com que a demanda seja maior do que a normal e a prevista. Esta informação deve ser manualmente incluída na previsão.

Conforme Kotler (1998), as previsões de vendas são usadas e têm impacto decisivo para muitos departamentos da empresa, como:

- Área Financeira, para dimensionar o dinheiro necessário aos investimentos e operações.
- Área de Manufatura, para definir os níveis de capacidade e produção de bens.
- Área Comercial, para adquirir o volume correto de suprimentos.
- Área de Recursos Humanos, para contratar o número de operários necessários, etc.

2.2.1. Tipos de Previsão

Os diversos métodos de previsão de demanda podem ser classificados em duas categorias principais: os métodos qualitativos e os métodos quantitativos (DAVIS et al., 2001).

Nesse item trataremos os dois principais métodos de previsão utilizados, apresentando seus benefícios, seus problemas e os casos em que deveriam ser utilizados.

2.2.1.1. Métodos Qualitativos

Os métodos qualitativos baseiam-se no julgamento e na experiência de pessoas que possam, por suas próprias características e conhecimentos, emitir opiniões sobre eventos futuros de interesse. Nesse sentido, o uso de julgamento pessoal não se restringe de forma alguma às previsões de demanda, podendo ser aplicado na análise de movimentos do comércio internacional, rumos da tecnologia, tendências de novos produtos, futuras condições econômicas e políticas, dentre outros (MOREIRA, 1996).

O método qualitativo é muito útil quando é realizado o lançamento de um novo produto para o mercado, em que não se tenha histórico de venda, assim como quando não se tem dados confiáveis sobre o histórico de venda dos produtos.

São denominados qualitativos os métodos que propiciam a incorporação do julgamento humano ao processo de previsão de vendas, mediante a interpretação sistemática de informações não capturáveis pelos modelos estatísticos, como intenções de compradores e opiniões de especialistas. As técnicas qualitativas aplicadas com mais sucesso são a pesquisa de mercado, o consenso de executivos, a composição da força de vendas e o método Delphi (KRAJEWSKI e RITZMAN, 1999).

A seguir será apresentado o funcionamento das técnicas citadas:

- Pesquisa de Mercado: a presença da pesquisa de mercado no método qualitativo deve-se ao simples fato de que os clientes determinam a demanda. Geralmente as pesquisas são feitas utilizando questionários e selecionando os clientes aleatoriamente, mas em alguns casos a pesquisa é dirigida para clientes em potencial. Segundo Moreira (1996), as pesquisas de mercado requerem

conhecimentos especializados e um planejamento bem estruturado para serem confiáveis.

- **Consenso de Executivos:** o consenso de executivos é um método baseado no conhecimento dos especialistas e profissionais com experiência na área; e por meio de discussões entre esse grupo é tomada uma decisão. O método é principalmente válido quando se trata de produtos novos, em que não se tem um histórico. Segundo Krajewski e Ritzman (1999), o consenso de executivos é frequentemente empregado para estimar o impacto provocado nas vendas por eventos promocionais ou outras circunstâncias especiais. As deficiências do método residem na dificuldade para obtenção de um consenso autêntico, quando conflitos de interesse emergem no grupo.
- **Composição da força de vendas:** desenvolver previsões com base na opinião do pessoal envolvido diretamente com as vendas pode ser uma alternativa atraente. Afinal, esses indivíduos mantêm contato diário tanto com os produtos da companhia quanto com os consumidores, conhecem o desenvolvimento histórico dos produtos e percebem as evoluções do mercado (MOREIRA, 1996). Esse método é muito interessante, utilizando o pessoal que trabalha direto com os clientes, pode trazer o benefício de se ter informação privilegiada, compreendendo melhor as tendências de mercado, assim como a visibilidade de um produto na praça. O maior problema dessa técnica é a subjetividade, podendo a informação trazida nem sempre condizer com a realidade.
- **Método Delphi:** esse consiste na reunião de um grupo de pessoas que deve opinar sobre determinado assunto, dentro das regras predeterminadas para a coleta e a depuração das opiniões. O método foi usado pela primeira vez em 1948 pela RAND Corporation, visando avaliar o impacto potencial de um ataque à bomba atômica sobre os Estados Unidos. A partir daí, sua aplicação estendeu-se a outras áreas, particularmente na previsão tecnológica, ou seja, na avaliação de mudanças na tecnologia existente e seu impacto sobre a organização. Envolve, geralmente, situações de longo prazo, nas quais os dados são escassos ou, mesmo, inexistentes, sendo o julgamento pessoal uma das poucas alternativas abertas à previsão. Aliás, o nome Delphi deriva de Delfos, o oráculo grego que supostamente tinha poderes de prever eventos futuros (MOREIRA, 1996). Em geral, o método consiste em fazer as pessoas integrantes do grupo responderem um questionário com as suas opiniões, é, então, feito um sumário com todas as opiniões coletadas e, logo, as

opiniões são apresentadas para todo o grupo, dando-se destaque às opiniões que divergem mais da opinião geral do grupo. Após reflexão sobre as opiniões apresentadas, é perguntado ao grupo se desejam refazer as suas opiniões, e, então, é refeito o processo, buscando que ao final, o alcance um consenso. Segundo Nahmias (1993), como qualquer outro método, o Delphi tem as suas vantagens e desvantagens. Sua vantagem é que ele permite a obtenção de opiniões pessoais sem que se estabeleçam interações dentro do grupo, as quais poderiam distorcer os resultados. Entretanto, é muito sensível à qualidade do instrumento de coleta de opiniões: como o contato pessoal é evitado, não há mecanismos para o debate de eventual ambigüidade em algumas questões.

2.2.1.2. Métodos Quantitativos

Os métodos quantitativos são aqueles que utilizam modelos matemáticos, com base na estatística, para análise dos dados, como forma de realizar a previsão, permitindo que diferentes especialistas obtenham previsões idênticas (MOREIRA, 1996).

A partir do momento em que as corporações, através de seus ERPs, sejam dotadas da capacidade de armazenar uma grande quantidade de informações – entre essas informações está incluso todo o histórico de venda de seus produtos – e utilizá-las para prever o comportamento de venda de seus produtos pode se tornar um grande diferencial competitivo.

Pode-se dizer que os métodos quantitativos se dividem em dois grupos, sendo eles:

- Métodos Casuais: os métodos casuais são métodos em que são relacionadas variáveis internas e externas sobre a empresa e a demanda de determinado produto. Essas variáveis, chamadas de variáveis causais, são determinadas a partir da ligação lógica entre as variáveis e o produto em questão. Se forem determinadas corretamente, é possível obter um resultado considerável na predição de demanda.
- Séries Temporais: a análise de séries temporais é um método que considera que os eventos ocorridos no passado representam significativamente o que virá ocorrer no futuro. Esse método é mais indicado para um pequeno espaço de tempo, devido à alterações no ambiente poderem afetar as previsões, quanto mais curto o tempo, mais precisa a previsão será. Segundo Moreira (1996), se o período coberto for suficientemente longo, o padrão de demanda resultante permite distinguir quatro

comportamentos ou efeitos associados a uma série temporal: I) Efeito de tendência: confere à demanda uma tendência a crescer ou a decrescer com o tempo. Pode ocorrer, também, de a demanda manter-se estacionária, variando sempre em torno de um valor médio; II) Efeito sazonal (ou estacional): representa o fato de que a demanda de muitas mercadorias assume comportamento semelhante em épocas bem definidas ao ano. Por exemplo, em um gráfico parte do seu aspecto de “dente de serra” pode ser atribuído ao efeito sazonal; III) Ciclo de negócios: constitui-se em flutuações econômicas de ordem geral, de periodicidade variável, decorrentes de uma multiplicidade de causas ainda em debate. São movimentos típicos da economia capitalista moderna, de difícil previsão; IV) Variações irregulares ou ao acaso: como o nome indica, são variáveis devidas a causas não identificadas, diferentemente dos ciclos de negócios. Pelo simples fato de ocorrerem ao acaso, tais variações não podem ser previstas por nenhum modelo de previsão.

No item a seguir será apresentado um método quantitativo chamado *Moving Average*, que é o método que a empresa estudada utiliza em suas previsões.

2.3.MOVING AVERAGE

Para uma análise e descrição preliminar, médias sumarizadas podem ser utilizadas para descrever uma série temporal. Alguns desses sumários estatísticos – como por exemplo, *simple average*, *single moving average*, *centered moving average*, ou possivelmente o *double moving average* – podem ser utilizados para suavizar uma série (YAFEE, 2000).

O *Single Moving Average* ou simplesmente *Moving Average* (MA) é um método simples que busca prever o próximo período por meio da utilização da média de um número especificado de períodos anteriores. Cada vez que o cálculo é efetuado, novos dados são utilizados, iniciando no dia em que o cálculo está sendo realizado, movendo sobre o tempo.

O método *Moving Average* pode ser definido por:

$$F_t = \frac{\sum_{i=1}^N S_{t-i}}{N}$$

Onde:

F_t = Previsão para o período t ,

S_{t-1} = Dados,

N = Número de Períodos.

O MA traz o benefício de suavizar as previsões realizadas a partir de uma série temporal com flutuações irregulares, assim como suavizar características como variações em função de sazonalidade inerente a cada problema.

Conforme será exposto na sequência do texto, foi constatado que a extensão da suavidade depende da ordem (quantidade de períodos) da série. Quanto maior for o período, maior deverá ser a suavização, mas em cenários instáveis, quanto maior o período maior poderá ser o ruído nos dados utilizados.

Na tabela abaixo é apresentado um exemplo de previsão de vendas utilizando o MA(3) (*Moving Average* com número de períodos igual a 3), sendo cada período equivalente a um mês; para prever cada mês, serão utilizados os 3 meses anteriores ao mês de previsão.

Tabela 2.1: Exemplo do MA de ordem 3.

Mês	Quantidade Vendida em Unidades	<i>Moving Average</i> de ordem 3
Janeiro	55	—
Fevereiro	43	—
Março	67	—
Abril	60	$(55 + 43 + 67) / 3 = 55,00$
Maio	70	$(43 + 67 + 60) / 3 = 56,66$
Junho	72	65,66
Julho	68	67,33
Agosto	63	70,00
Setembro	93	67,66
Outubro	80	74,66
Novembro	74	78,66
Dezembro	68	82,33

O MA de ordem 1 (utilizando apenas um mês), conhecido também como *naive forecasting*, utiliza o último período como previsão para o período seguinte (YAFEE, 2000).

Esse capítulo teve como objetivo apresentar o cenário onde o problema está inserido, mostrando os sistemas de informação de onde as informações foram coletadas, os tipos de previsão de vendas, assim como o método quantitativo utilizado pela empresa estudada.

No capítulo a seguir serão apresentadas as Redes Neurais Artificiais, que é o método utilizado por esta dissertação, como uma alternativa aos *Forecastings* clássicos, apresentados neste capítulo.

3. REDES NEURAIIS ARTIFICIAIS

Nesse capítulo serão apresentadas as Redes Neurais Artificiais, método proposto para a solução do problema de previsão de vendas. Nos itens a seguir serão apresentados a definição de uma RNA, o funcionamento do Neurônio Biológico, o *Perceptron*, o *Perceptron* de Múltipla Camada (incluindo seu treinamento), assim como os benefícios na utilização de RNAs. Para fins de conclusão deste capítulo serão apresentados trabalhos correlatos de RNAs aplicadas a problemas de predição.

3.1. DEFINIÇÃO

O termo Rede Neural Artificial, raro algumas décadas atrás na literatura científica, representa hoje uma vigorosa área de pesquisa multidisciplinar. É frequentemente identificada como uma subespecialidade da Inteligência Artificial, outras vezes como uma classe de modelos matemáticos para problemas de classificação e reconhecimento de padrões, outras ainda como uma parte da teoria conexionista dos processos mentais e finalmente, como uma categoria de modelos em ciência da cognição (KOVÁCS, 1996).

Segundo Haykin (2001), as Redes Neurais, ou Redes Neurais Artificiais, representam uma tecnologia que tem raízes em muitas disciplinas: neurociência, matemática, estatística, física, ciência da computação e engenharia. As redes neurais encontram aplicações em campos bastante diversos, sendo que entre as principais tarefas a que se destinam estão: a modelagem, a análise de séries temporais, o reconhecimento de padrões, o processamento de sinais e controles.

Nas seções a seguir serão apresentados conceitos fundamentais que servem de embasamento do funcionamento de uma RNA. Dentre os itens destacados serão apresentados: o Neurônio Biológico, o Neurônio Artificial, os tipos de RNAs, o *Perceptron*, o *Perceptron* Multicamada, assim como o processo de treinamento das Redes Neurais Artificiais Multicamadas.

3.2.O NEURÔNIO BIOLÓGICO

O neurônio é a unidade fundamental para o funcionamento de todos os tecidos do sistema nervoso, incluindo o cérebro. Cada neurônio consiste em um corpo celular, ou soma, que contém um núcleo celular. As ramificações a partir do corpo celular são um número de fibras chamadas de dendritos e uma longa e única fibra chamada de axônio. Os dendritos ramificam-se em uma densa rede em volta da célula, enquanto o axônio se estende por uma longa distância, normalmente em torno de um centímetro (100 vezes o diâmetro do corpo celular) e tão longo quanto um metro em casos extremos. Eventualmente, o axônio também se ramifica em vertentes e sub-vertentes que se ligam aos dendritos e corpos celulares de outros neurônios. Essa conexão é chamada de Sinapse. Cada neurônio forma sinapses de qualquer lugar com dúzias até centenas de milhares de outros neurônios (RUSSELL e NORVIG, 1995). Abaixo é ilustrada as partes que compõem um neurônio:

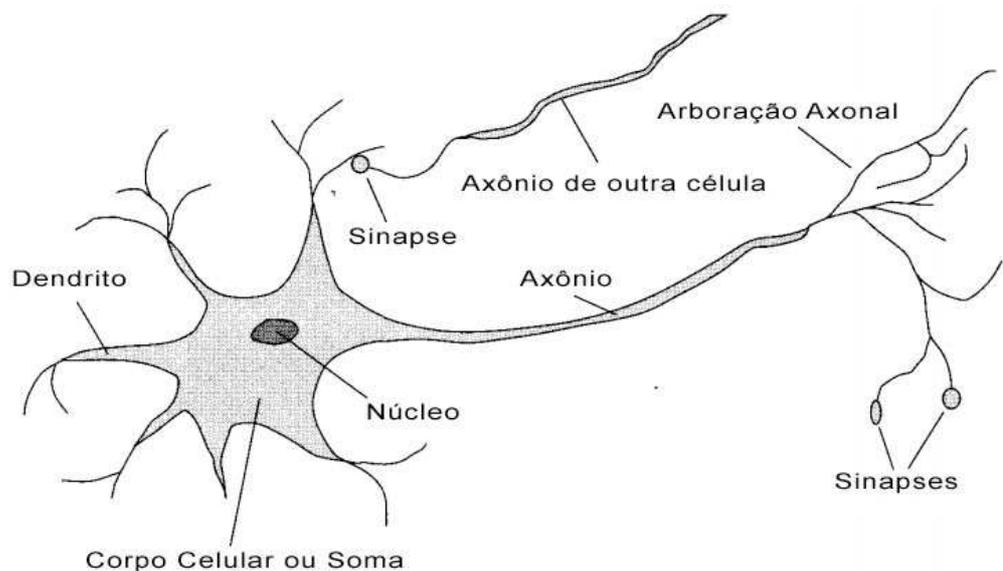


Figura 3.1 As partes de um neurônio, adaptado de Russell e Norvig (1995).

As manifestações elétricas de neurônios biológicos foram observadas pela primeira vez no século 19 por DuBois Reymond com o auxílio de galvanômetros, que é um instrumento utilizado para medir energia elétrica. O funcionamento dessas células começou a ser melhor entendido com a invenção, no final do século 19, por Crookes, do tubo de raios catódicos e com a sua aplicação à observação da atividade elétrica nervosa, principalmente por Erlanger e Gasser. Essas investigações lhes renderam o prêmio Nobel de fisiologia em 1944. Nas duas décadas seguintes, com o resultado do trabalho de vários pesquisadores,

passou-se a entender o neurônio biológico como sendo basicamente o dispositivo computacional elementar do sistema nervoso, que possuía entradas (muitas entradas) e saída. As entradas ocorrem através das conexões sinápticas, que conectam a árvore dendrital aos axônios de outras células nervosas. Os sinais que chegam por esses axônios são pulsos elétricos conhecidos como impulsos nervosos ou potenciais de ação, e constituem a informação que o neurônio processará, de alguma forma, para produzir como saída um impulso nervoso no seu axônio (KOVÁCS, 1996).

Segundo Kovács (1996), pode-se dizer que o neurônio opera como um dispositivo que codifica em frequência de pulsos o resultado da integração espacial/temporal dos estímulos. No diagrama da figura abaixo é representada uma relação típica entre a frequência de impulsos nervosos e a intensidade do estímulo depolarizante.

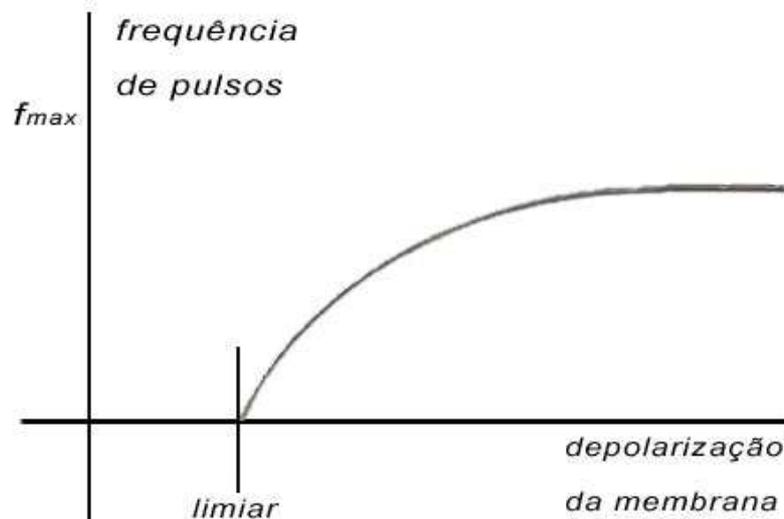


Figura 3.2: Relação entre a frequência de pulsos de um neurônio e a intensidade do estímulo de longa duração.

No item a seguir será apresentado o funcionamento de um neurônio artificial.

3.3.NEURÔNIO ARTIFICIAL

A unidade básica de uma RNA é o neurônio artificial, o qual teve a primeira aparição reconhecida em 1943 através da publicação do artigo *A Logical Calculus of the Ideas Immanent in Nervous Activity* por McCulloch e Pitts. Eles propuseram um modelo de neurônios artificiais no qual cada neurônio é caracterizado por ter o estado de ligado ou desligado, com uma chave para ligado ocorrendo em resposta a um estímulo por um número

adequado de neurônios vizinhos, sendo que naquela época especulava-se que a natureza da inteligência era booleana (RUSSELL e NORVIG, 1995). Após mais uma década de descrença nas Redes Neurais, Rumelhart, Hinton e Williams (1986) propuseram um modelo não linear para o neurônio biológico, o qual é amplamente aceito pela comunidade científica.

O neurônio artificial ou unidade processamento é apresentado na figura abaixo:

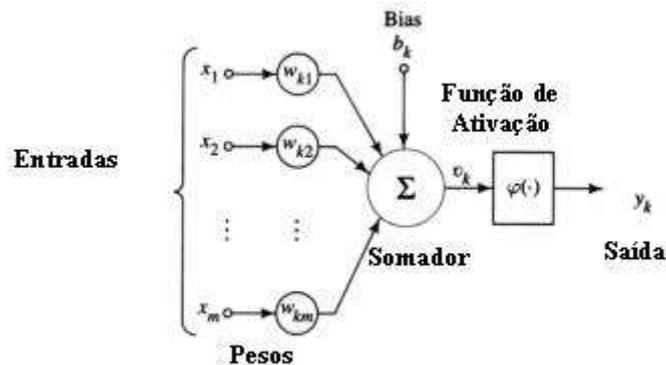


Figura 3.3: Neurônio Artificial ou Unidade de Processamento.

Conforme demonstrado na figura acima, o neurônio artificial é composto basicamente por entradas, onde cada entrada x_i é multiplicada por um peso sináptico w_k e então todas são somadas juntamente com o *bias*, que é um parâmetro externo de equilíbrio do neurônio artificial conectado a uma entrada com valor fixo, por final aplicando a função de ativação φ , gerando o valor de saída do neurônio artificial.

Se realizada uma analogia entre o neurônio biológico e o artificial, pode-se dizer que os dendritos foram substituídos pelas entradas, a sinapse foi substituída pelos pesos sinápticos e a função de ativação representa a modelagem da frequência de pulsos de um neurônio biológico em relação a intensidade do estímulo de longa duração.

As funções de ativação mais populares são as funções sigmóide, tangente hiperbólica, reta e degrau. Na tabela abaixo serão apresentadas as funções de ativação, assim como as suas respectivas derivadas, que serão utilizadas na aplicação do algoritmo *Backpropagation*, apresentado no decorrer do capítulo.

Tabela 3.1: Funções de Ativação mais utilizadas.

Função	Equação	Derivada
Sigmóide	$g(v) = (1 + \exp[-\beta v])^{-1}$	$g'(v) = \beta \cdot (1 - g(v)) \cdot g(v)$
Tangente Hiperbólica	$g(v) = \frac{1 - \exp(-2v)}{1 + \exp(2v)}$	$g'(v) = \beta \cdot (1 - g(v))^2$
Reta	$g(v) = \alpha \cdot v$	$g'(v) = \alpha$
Degrau	$g(v) = h(v)$	-----

Segundo Russell e Norvig (1995), a RNA é caracterizada por conter conexões entre diversos neurônios, cada conexão entre os neurônios contém um peso sináptico associado, onde o conhecimento fica armazenado. Os neurônios são arranjados em camadas, sendo que as camadas de entrada e de saída são ligadas ao mundo externo, podendo conter uma ou mais camadas intermediárias entre elas. O arranjo da RNA, como a definição do número de camadas, número de neurônio em cada camada, função de ativação, é denominado arquitetura da RNA.

Uma RNA com todas as entradas conectadas diretamente na saída são chamadas redes de camada única ou também conhecida como rede Perceptron. Já as redes multicamadas ou Perceptron de Múltiplas Camadas são as redes que contém uma ou mais camadas escondidas entre as entradas e as saídas. As redes Perceptron e as redes Perceptron de Múltiplas camadas serão apresentadas nos tópicos a seguir.

3.4.O PERCEPTRON

Segundo Haykin (2001), o *Perceptron* é a forma mais simples de uma rede neural usada para a classificação de padrões linearmente separáveis. Basicamente, ele consiste em um único neurônio com pesos sinápticos ajustáveis e bias.

O estudo dos *Perceptrons* iniciou-se nos anos 50. Embora redes de todos os tamanhos e topologias tenham sido consideradas, o único método de aprendizado efetivo naquele momento foi o das redes de camada única, por isso nesta a maioria dos esforços foram gastos. Hoje o nome *Perceptron* é usado como um sinônimo para redes de camada única (RUSSELL e NORVIG, 1995).

O *Perceptron* ganhou proporções maiores através de Rosenblatt (1962), que utilizou o princípio de aprendizado de Hebb (1949). Segundo este, o aprendizado ocorria por meio de exemplos.

O problema que Rosenblatt propôs resolver foi o de casos simples como a implementação das funções booleanas E e OU de duas variáveis, em que é relativamente trivial escolher os ganhos sinápticos e o valor do limiar. Porém, para a implementação de uma função discriminatória arbitrária esta escolha é não trivial e, dependendo do número de variáveis envolvidas, sem a existência de algum método, beira o impossível (KOVÁCS, 1996).

Na figura abaixo é representado uma rede *Perceptron* com três entradas e uma saída:

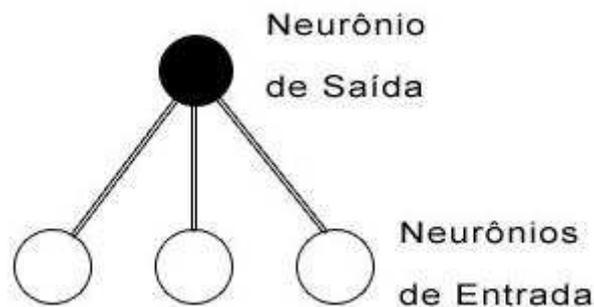


Figura 3.4: Exemplo de Rede *Perceptron*

Em 1969, Minsky e Papert criticaram o *Perceptron* demonstrando que este com uma camada simples, não poderia simular o comportamento de uma função XOR (ou-exclusivo). Os anos 60 e 70 são considerados os anos negros das redes neurais, um período em que ocorreram poucas pesquisas na área devido a sua limitação.

Na década de 80, Rumelhart, Hinton e Williams desenvolveram um algoritmo conhecido como *Backpropagation*. Esse algoritmo é conhecido como o principal responsável pelo retorno dos estudos sobre as RNAs. Através dele foram criadas as redes *Perceptron* de Múltiplas camadas que, quando treinadas adequadamente, têm um grande poder de generalização podendo aproximar qualquer função contínua. Os *Perceptrons* de Múltiplas camadas serão apresentados no item a seguir.

3.5. PERCEPTRON DE MÚLTIPLAS CAMADAS

Nessa seção será apresentada uma das mais importantes arquiteturas de redes neurais, as redes *Perceptron* de Múltiplas Camadas. Esta é uma rede do tipo *Perceptron* com pelo menos uma camada intermediária.

Os *Perceptrons* de Múltiplas Camadas têm sido aplicados com sucesso para resolver diversos problemas não linearmente separáveis, através do seu treinamento de forma supervisionada com o algoritmo de retropropagação do erro (*Backpropagation*) (HAYKIN, 2001). Basicamente, a aprendizagem por retropropagação de erro consiste em dois passos através das diferentes camadas da rede: um passo para frente, a propagação, e um passo para trás, a retropropagação. Na propagação, um padrão de atividade (vetor de entrada) é aplicado aos nós sensoriais da rede e seu efeito se propaga através da rede, camada por camada. Finalmente, um conjunto de saídas é produzido como a resposta real da rede. Durante o passo de propagação, os pesos sinápticos da rede são todos fixos. Durante a retropropagação, por outro lado, os pesos sinápticos são todos ajustados de acordo com uma regra de correção de erro. Especificamente, a resposta real da rede é subtraída de uma resposta desejada (alvo) para produzir um sinal de erro. Este sinal de erro é então propagado para trás por intermédio da rede, contra a direção das conexões sinápticas – vindo daí o nome de “retropropagação do erro” (*Backpropagation*). Os pesos sinápticos são ajustados para fazer com que a resposta real da rede se mova para mais perto da resposta desejada, em um sentido estatístico (HAYKIN, 2001).

No item a seguir será apresentada a arquitetura de uma RNA *Perceptron* de Múltiplas Camadas.

3.5.1. Arquitetura dos *Perceptrons* de Múltiplas Camadas

A arquitetura das redes *Multilayer Perceptron* (MLP) ou *Perceptron* de Múltiplas Camadas é dada pela camada de entrada (*input layer*), por uma ou mais camadas escondidas (*hidden layer*), e pela camada de saída (*output layer*). Tanto a camada de entrada como a camada de saída geralmente são traçadas conforme as variáveis do problema. Já a camada

escondida é dada por característica do projeto, por meio de simulações, identificando qual seria a melhor arquitetura para resolver o problema em questão.

Abaixo é apresentado um exemplo de rede neural *Perceptron* de Múltiplas Camadas:

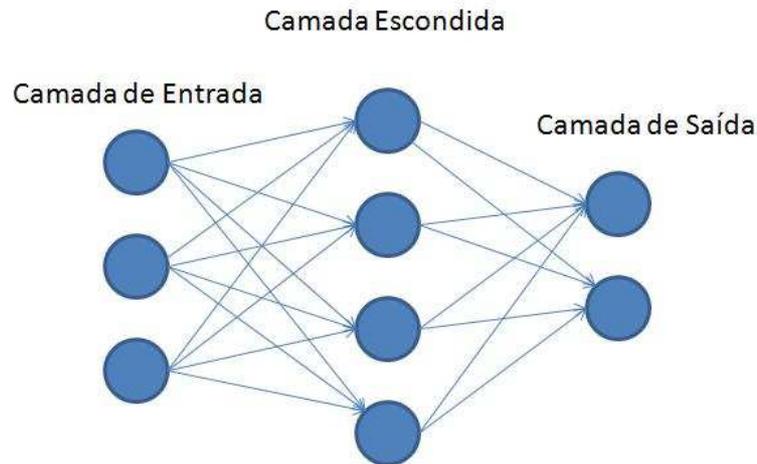


Figura 3.5: Rede Neural MLP com uma camada escondida.

Quanto maior for a quantidade de neurônios na camada escondida, maior será a habilidade da rede de lidar com informações não lineares. No entanto, deve-se ter cuidado no momento de estruturar a rede, pois se houver neurônios em excesso na camada escondida, é possível que a rede informe um erro baixo como resposta, entretanto não esteja apta a realizar previsões com dados não utilizados em seu treinamento, fazendo com que a rede “decore” os valores utilizados, perdendo o poder de generalização, conhecido como *overfitting*. Já o oposto, isto é, poucos neurônios na camada escondida, pode ocasionar o que é conhecido por *underfitting*, que significa a rede ter um erro alto, não sendo capaz de aproximar os resultados, mesmo para os dados utilizados em seu treinamento.

No item a seguir será apresentado o treinamento de uma RNA assim como o Algoritmo *Backpropagation*, que é o mais utilizado em redes *Perceptron* de Múltiplas Camadas, também utilizado neste estudo.

3.6.TREINAMENTO DE REDES NEURAIAS ARTIFICIAIS

O objetivo principal do treinamento é atribuir valores aos pesos sinápticos, que são iniciados de forma aleatória. Uma vez apresentados os dados à rede, também conhecidos como

exemplos, é então verificada como a RNA responde a essas entradas. Caso a resposta da rede comparada com a resposta esperada não seja satisfatória, os pesos devem ser ajustados pelo algoritmo selecionado de forma a reduzir o erro da RNA. Esse processo deve ser repetido até que algum critério de parada pré-estabelecido seja atingido (WU e MCLARTY, 2000).

É por intermédio dos treinamentos que se encontra qual a melhor arquitetura da RNA, ou seja, qual é a quantidade de camadas e de neurônios em cada camada ideal para o problema. Essa identificação é realizada através da alteração dos parâmetros em cada treinamento, buscando selecionar a estrutura que gerou o menor erro.

O processo de aprendizagem da rede, o qual ocorre no treinamento através do ajuste dos pesos, é realizado por um algoritmo de aprendizagem específico. O objetivo de se efetuar o treinamento da rede é a minimização de uma métrica de erro no treinamento desta, buscando uma aproximação numérica entre a saída obtida e o que se deseja que ela produza, pois é desejado que rede atinja um erro global mínimo, que representa a generalização do aprendizado. Entretanto, é constante a presença de mínimos locais devido a questões relacionadas a ruídos nos dados de entrada e a configuração dos parâmetros de treinamento do modelo, que tornam árdua a tarefa de treinamento. Os métodos de treinamento podem ser divididos em: supervisionados e não-supervisionados.

Na aprendizagem supervisionada, que é também denominada aprendizagem mediada por um professor, basicamente são apresentadas entradas e suas respectivas saídas desejadas, de modo a se ajustarem os parâmetros da rede sobre a influência do sinal de erro. O sinal de erro é obtido por meio da diferença entre a resposta desejada e a resposta obtida pela saída da rede. O treinamento das redes MLP se enquadra nesse tipo de aprendizagem. Já a aprendizagem não-supervisionada, também conhecida como aprendizagem sem professor, não há exemplos rotulados da função a ser aprendida (mapeamento entrada-saída) (HAYKIN, 2001).

Os algoritmos de aprendizagem apresentam uma constante conhecida como taxa de aprendizagem, que controla a velocidade e a qualidade do ajuste dos pesos da rede. Caso o valor desta constante seja baixo, os pesos se atualizam de forma suave, mas também de forma lenta. Contudo, uma taxa de aprendizagem alta faz o ajuste dos pesos sinápticos ser mais rápido, podendo a rede se tornar instável, oscilando os valores sem chegar no erro mínimo desejado.

O algoritmo de aprendizagem ainda pode contar com uma variável chamada constante de momento. Esta faz uma relação entre as mudanças anteriores dos pesos na direção do movimento atual no espaço de pesos, buscando impedir que o treinamento caia em mínimos

locais, que é uma área da função onde o erro ainda está longe do menor possível, chamado de mínimo global.

A seguir será apresentado o algoritmo Backpropagation, algoritmo mais utilizado no treinamento de redes MLP.

3.6.1.O Algoritmo Backpropagation

O algoritmo *Backpropagation* ou da retropropagação do erro, para redes de neurônios de múltiplas camadas, foi essencialmente inventado e popularizado por Rumelhart, Hinton e Williams, resolvendo uma das limitações fundamentais para o treinamento de redes complexas, que é a capacidade de lidar com dados não linearmente separáveis (KOVÁCS, 1996).

O algoritmo *Backpropagation* tem por objetivo utilizar a retropropagação do erro buscando ajustar os pesos sinápticos de modo a reduzir o erro da rede.

O algoritmo é, em essência, definido por duas etapas: I) a propagação, em que após apresentado dados para a camada de entrada da rede, os valores são propagados pela rede, camada por camada, até chegar na camada de saída, onde então, por meio da comparação entre a resposta obtida e a resposta desejada, é obtido o erro; II) retropropagação, em que desde a camada de saída até a camada de entrada, são ajustados os pesos sinápticos de forma a diminuir o erro do treinamento.

3.6.1.1. Treinamento com *Backpropagation*

Aqui serão apresentadas as etapas que compõem o treinamento de uma rede *Perceptron* de Múltipla Camada, sendo elas: a inicialização, a apresentação dos exemplos de treinamento, a computação para frente (Propagação), o cálculo do erro e a computação para trás (Retropropagação) (HAYKIN, 2001).

1. Inicialização – assumindo que nenhuma informação prévia esteja disponível, retire os pesos sinápticos e limiares de uma distribuição uniforme cuja média é zero e

cuja variância é escolhida para que o desvio padrão dos campos locais induzidos dos neurônios se encontre na transição entre as partes linear e saturada da função de ativação.

2. Apresentação dos Exemplos de Treinamento – apresentar um exemplo do conjunto de treinamento para a RNA e proceder com os itens 3 e 4.
3. Computação para Frente (Propagação) – a computação para frente ocorre da seguinte forma: Suponha que um exemplo de treinamento de época seja representado por $(\mathbf{x}(n), \mathbf{d}(n))$, com o vetor de entrada $\mathbf{x}(n)$ aplicado à camada de entrada de nós (neurônios) sensoriais e vetor de resposta desejada $\mathbf{d}(n)$ apresentado à camada de saída de nós computacionais. Calcule os campos locais induzidos e os sinais funcionais da rede prosseguindo para frente através da rede, camada por camada. O campo local induzido $v_j^{(l)}(n)$ para o neurônio j na camada l é

$$v_j^{(l)} = \sum_{i=0}^{m_a} w_{ji}^{(l)}(n) y_i^{(l-1)}(n),$$

Em que $y_i^{(l-1)}(n)$ é o sinal (função) de saída do neurônio i na camada anterior $l-1$, na interação n , e $w_{ji}^{(l)}(n)$ é o peso sináptico do neurônio j na camada l , que é alimentado pelo neurônio i da camada $l-1$. Para $i=0$, temos que $y_0^{(l-1)}(n) = +1$ e $w_{j0}^{(l)} = b_j^{(l)}(n)$ é o bias aplicado ao neurônio j na camada l . Assumindo-se o uso de uma função sigmóide, o sinal de saída do neurônio j na camada l é

$$y_j^{(l)} = \varphi_j(v_j^{(l)}(n)),$$

se o neurônio j está na primeira camada oculta ($l=1$), faça

$$y_j^{(1)}(n) = x_j(n),$$

em que $x_j(n)$ é o j -ésimo elemento do vetor de entrada $\mathbf{x}(n)$. Se o neurônio j está na camada de saída ($l=L$, onde L é denominado a profundidade da rede), faça

$$y_j^{(L)} = o_j(n)$$

4. Calcule o sinal do erro: o cálculo é dado por

$$e_j(n) = d_j(n) - o_j(n)$$

em que $d_j(n)$ é o j -ésimo elemento do vetor resposta desejada $\mathbf{d}(n)$.

5. Computação para trás (Retropropagação) – a computação para trás é dada da seguinte forma: Calcule os δ s (gradientes locais) da rede, definidos por

$$\delta_j^{(l)}(n) = \begin{cases} e_j^{(l)}(n) \varphi_j'(v_j^{(l)}(n)), & \text{neurônio } j \text{ da camada de saída } L, \\ \varphi_j'(v_j^{(l)}(n))' \sum_k \delta_k^{(l+1)}(n) w_{kj}^{(l+1)}(n), & \text{neurônio } j \text{ na camada oculta } l, \end{cases}$$

sendo que o apóstrofe em $\varphi_j'(x)$ representa a diferenciação em relação ao argumento. Ajuste os pesos sinápticos da rede na camada l de acordo com a regra delta generalizada:

$$w_{ji}^{(l)}(n+1) = w_{ji}^{(l)}(n) + \alpha [w_{ji}^{(l)}(n-1)] + \mu \delta_j^{(l)}(n) y_i^{(l-1)}(n)$$

sendo que μ é o parâmetro da taxa de aprendizagem e α é a constante de momento.

6. Iteração – refazer as etapas 3, 4 e 5 com os outros exemplos do treinamento até que a condição de parada seja satisfeita, isto é, o erro da rede estar em um patamar aceitável conforme o problema, ou atingir um número de ciclos iterativos de treinamento pré-fixados.

3.7. PROPRIEDADES E CAPACIDADES DAS RNAS

Esta seção tem o intuito de apresentar as propriedades e capacidades das RNAs, visando a compreensão de alguns dos motivos pelos quais fazem elas serem amplamente utilizadas, assim como permitir a abstração de ambientes que se beneficiariam do seu uso.

Conforme Haykin (2001), o uso de Redes Neurais Artificiais oferece propriedades úteis e capacidades que serão apresentadas a seguir:

- Não-linearidade: Um neurônio artificial pode ser linear ou não-linear. Uma rede neural, constituída por conexões de neurônios não-lineares é ela mesma não-linear. Além disso, a não-linearidade é de um tipo muito especial, no sentido de ser distribuída por toda a rede. A não-linearidade é uma propriedade muito importante, particularmente se o mecanismo físico responsável pela geração do sinal de entrada (p. ex., sinal de voz) for inerentemente não-linear.
- Mapeamento de Entrada-Saída: Um paradigma popular de aprendizagem chamado aprendizagem mediada por um professor ou aprendizagem supervisionada envolve

a modificação dos pesos sinápticos de uma rede neural para aplicação de um conjunto de amostras de treinamento rotuladas ou exemplos da tarefa. Cada exemplo consiste em um sinal de entrada único e em uma resposta desejada correspondente. Apresenta-se para a rede um exemplo escolhido ao acaso do conjunto, e os pesos sinápticos (parâmetros livres) da rede são modificados para minimizar a diferença entre a resposta desejada e a resposta real da rede, produzida pelo sinal de entrada, de acordo com um critério estatístico apropriado. O treinamento da rede é repetido para muitos exemplos do conjunto até que a rede alcance um estado estável em que não haja mais modificações significativas nos pesos sinápticos. Os exemplos de treinamento, previamente aplicados, podem ser reaplicados durante a sessão de treinamento, mas em uma ordem diferente. Assim, a rede aprende através dos exemplos ao construir um mapeamento de entrada-saída para o problema considerado. Tal abordagem nos faz lembrar do estudo de inferência estatística não-paramétrica, que é um ramo da estatística que trata da estimação independente do modelo ou, do ponto de vista biológico, aprendizagem tabula rasa; o termo “não-paramétrico” é utilizado aqui para significar o fato de que não são feitas suposições prévias sobre o modelo estatístico dos dados de entrada. Considere, por exemplo, uma tarefa de classificação de padrões, cujo o objetivo seja atribuir um sinal de entrada representando um objeto físico ou evento a uma dentre várias categorias (classes) preestabelecidas. Em uma abordagem não-paramétrica, para esse problema, o objetivo é “estimar” fronteiras de decisão arbitrárias no espaço do sinal de entrada para a tarefa de classificação de padrões utilizando um conjunto de exemplos, e fazê-lo sem invocar um modelo de distribuição probabilístico. Um ponto de vista similar está implícito no paradigma de aprendizagem supervisionada, o que sugere uma analogia próxima entre o mapeamento de entrada-saída realizado por uma rede neural e a inferência estatística não-paramétrica.

- Adaptabilidade: As redes neurais tem uma capacidade inata de adaptar seus pesos sinápticos à modificações do meio ambiente. Em particular, uma rede neural treinada para operar em um ambiente específico pode ser facilmente retreinada para lidar com pequenas modificações nas condições operativas do ambiente. Além disso, quando está operando em um ambiente não-estacionário (i.e., onde as estatísticas mudam com o tempo), uma rede neural pode ser projetada para modificar seus pesos sinápticos em tempo real. A arquitetura natural de uma rede

neural para classificação de padrões, processamento de sinais e aplicações de controle, aliada à capacidade de adaptação da rede, a torna uma ferramenta muito útil para classificação adaptativa de padrões, processamento adaptativo de sinais e controle adaptativo. Como regra geral, pode-se dizer que quanto mais adaptativo se fizer um sistema, assegurando-se que este se mantenha estável, mais robusto tenderá ser o seu desempenho quando for exigido operar em um ambiente não-estacionário. Contudo, deve ser enfatizado, que adaptabilidade nem sempre resulta em robustez; na verdade pode resultar no contrário. Um sistema adaptativo com constantes de tempo pequenas, por exemplo, pode se modificar rapidamente e, assim, tender a responder a perturbações espúrias, causando uma drástica degradação no desempenho do sistema.

- **Resposta a Evidências:** No contexto de classificação de padrões, uma rede neural pode ser projetada para fornecer informação não somente sobre qual padrão particular selecionar, mas também sobre a confiança ou crença na decisão tomada. Esta última informação pode ser utilizada para rejeitar padrões ambíguos, caso eles estejam presentes, e com isso melhorar o desempenho de classificação da rede.
- **Informação Contextual:** O conhecimento é representado pela própria estrutura e estado de ativação de uma rede neural. Cada neurônio da rede é potencialmente afetado pela atividade de todos os outros neurônios da rede. Conseqüentemente, a informação contextual é tratada naturalmente pela rede neural.
- **Tolerância a falhas:** Uma rede neural, implementada na forma física (em hardware), tem o potencial de ser inerentemente tolerante a falhas, ou capaz de realizar computação robusta, no sentido de que seu desempenho se degrada suavemente sob condições de operação adversas. Se um neurônio ou suas conexões são danificados, por exemplo, a recuperação de um padrão armazenada na rede, o dano deve ser extenso para que a resposta global da rede seja degradada seriamente. Assim, em princípio, uma rede neural exhibe uma degradação suave do desempenho em vez de apresentar uma falha catastrófica. Há algumas evidências empíricas para a computação robusta, mas geralmente ela não é controlada. Para se assegurar que uma rede neural seja de fato tolerante a falhas pode ser necessário adotar-se medidas corretivas no projeto do algoritmo utilizado para treinar a rede.
- **Implementação em very-large-scale-integration (VLSI):** A natureza maciçamente paralela de uma rede neural a faz ser potencialmente rápida na computação de

certas tarefas. Esta mesma característica torna uma rede neural adequada para implementação utilizando tecnologia de integração em escala muito ampla. Uma virtude benéfica particular da tecnologia VLSI é que ela fornece um meio de capturar comportamentos realmente complexos de uma forma altamente hierárquica.

- **Uniformidade de Análise e Projeto:** Basicamente, as redes neurais desfrutam de universalidade como processadores de informação. Isso no sentido de que a mesma notação é utilizada em todos os domínios envolvendo a aplicação de redes neurais. Esta característica se manifesta de diferentes modos: I) Os neurônios, de uma forma ou de outra, representam um ingrediente comum a todas as redes neurais; II) Essa uniformidade torna possível compartilhar teorias e algoritmos de aprendizagem em aplicações diferentes de redes neurais; III) Redes modulares podem ser construídas por meio de uma integração homogênea de módulos.
- **Analogia Neurobiológica:** O projeto de uma rede neural é motivado pela analogia ao cérebro, que é uma prova viva de que o processamento paralelo tolerante a falhas é não somente possível fisicamente mas também rápido e poderoso. Os neurobiólogos olham para as redes neurais (artificiais) como uma ferramenta de pesquisa para a interpretação de fenômenos neurobiológicos. Por outro lado, os engenheiros olham para a neurobiologia procurando novas ideias para resolver problemas mais complexos que aqueles baseados em técnicas convencionais de projeto por conexões fixas.

3.8. TRABALHOS CORRELATOS

Com o intuito de justificar a metodologia utilizando RNAs, esta seção tem o objetivo de apresentar trabalhos correlatos que utilizaram essa metodologia para soluções próximas ao problema abordado nesta dissertação. Devido à dificuldade de se encontrar trabalhos utilizando RNAs para previsão de vendas em indústrias, os trabalhos a seguir foram selecionados por apresentarem propostas de aplicação de RNAs ao problema de predição de valores futuros, só que em outras áreas.

3.8.1. Previsão de Demanda de Energia Elétrica

Previsão de energia elétrica tem sido foco de pesquisadores, contendo grande quantidade de trabalhos publicados nessa área. O objetivo de prever energia elétrica vem da necessidade das companhias de produção e distribuição de energia elétrica em organizar, da melhor forma possível, os seus recursos, assim, possibilitando um trabalho preventivo em uma das áreas vitais para a sociedade.

Entre as recentes publicações nessa área, podemos destacar *Artificial Neural Network Approach for Short Term Load Forecasting for Illam Region* (SHIRVANY e HAYATI, 2007), *A Multi-Layer Artificial Neural Network Architecture Design for Load Forecasting in Power Systems* (ARDIL, 2007), *An Optimized Adaptive Neural Network for Annual Midterm Energy Forecasting* (TSEKOURAS, 2006), *Application of Neural Networks for Short-Term Load Forecasting* (AFKHAMI e YAZDI, 2006), *Accurate Electricity Load Forecasting with Artificial Neural Networks* (ORTIZ-ARROYO, 2005) e *Forecasting of the electric energy demand trend and monthly fluctuation* (GONZÁLES-ROMERA, JARAMILLO-MORÁN e CARMONA-FERNÁNDEZ, 2007).

O trabalho *Forecasting of the electric energy demand trend and monthly fluctuation* será detalhado a seguir a fim de apresentar o cenário e as soluções encontradas, bem como os resultados obtidos. Este trabalho foi escolhido por apresentar aplicação semelhante ao realizado por esta dissertação, utilizando o método de previsão quantitativo *Moving Average* como fonte de informação para as RNAs.

3.8.1.1. O Problema da Previsão de Energia Elétrica

Segundo González-Romera, Jaramillo-Morán e Carmona-Fernández (2007), é muito importante prever a demanda de energia elétrica, tanto para quem produz como para quem distribui, separando elas em previsão de curto prazo e previsão de longo prazo. O trabalho tem por objetivo a realização de previsões de demandas mensais, justificando a necessidade desse

tipo de previsão para que possam ser feitos planos de manutenção e também para que sirvam como uma análise de mercado útil em vários aspectos, tanto para quem produz quanto para quem distribui energia. Previsões precisas são desejadas por apresentarem tendências que influenciam a evolução econômica e tecnológica do mercado de energia elétrica.

Os autores destacam que nesse cenário existe uma flutuação causada por questões sazonais, com isso, justificou a utilização de duas previsões, uma prevendo tendências no mercado de energia elétrica e outra prevendo as flutuações recorrentes do período de consumo analisado, assim, depois de realizadas as previsões, as mesmas são sumarizadas prevendo o consumo para o período desejado.

3.8.1.2. Estrutura das RNAs

Os autores utilizaram o *Multi-layer Perceptron* para realizar a previsão, por ela ser amplamente utilizada devido a sua habilidade de identificar a evolução temporal de um sistema dinâmico (CYBENCO, 1989).

Os dados utilizados como entrada da rede foram os consumos mensais de 12 períodos anteriores a data de previsão. Foi aplicado o método *Moving Average* antes de realizar a entrada dos dados na rede, buscando utilizar tendências de consumo no período como entrada da rede, suavizando possíveis alternâncias de consumo dentro do período. O período de 12 meses foi escolhido por meio de testes buscando identificar a melhor arquitetura.

Foi utilizada uma camada escondida com 2 neurônios, escolhida a partir de testes com a RNA utilizando uma função de avaliação dos resultados. A função de ativação utilizada foi a tangente hiperbólica, que trabalha com uma faixa de valores entre -1 e 1, devido a isso, os dados de entrada da rede foram normalizados de forma a ficarem dentro dos valores possíveis da função tangente hiperbólica.

A saída da rede é a previsão de consumo para o período seguinte.

3.8.1.3. Resultados

O trabalho apresentou resultados significantes, chegando a um erro médio de 2%, sendo que em raros casos o erro ficou acima de 5% e em todos os casos o erro ficou abaixo de 7%, apresentando as RNAs como uma excelente alternativa para as previsões no cenário em que o trabalho foi realizado.

3.8.2.Previsão de Vendas

Previsão de vendas, independente da área atuação, é uma das principais necessidades das empresas, em que a partir de uma correta previsão, é possível que a empresa possa organizar seus esforços de modo a reduzir seus custos e ao mesmo tempo garantir a qualidade que os clientes esperam.

Nesse cenário existem muitos esforços. Entre os trabalhos que utilizaram RNAs para previsão de vendas podemos citar os artigos *Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods* (ALON, QI e SADOWSKI, 2001), *Time series sales forecasting for short shelf-life food products based on artificial neural networks and evolutionary computing* (DOGANIS, ALEXANDRIDIS, PATRINOS e SARIMVEIS, 2006), *A hybrid econometric-neural network modeling approach for sales forecasting* (LUXHOJ, RIIS, e STENSBALLE, 1996), *Sales Forecasting Using Neural Networks* (THIESING e VORNBERGER, 1997) e *A sales forecasting system based on fuzzy neural network with initial weights generated by genetic algorithm* (KUO, 2001).

Dentre os artigos destacados são detalhados, abaixo: *A hybrid econometric-neural network modeling approach for sales forecasting* e o artigo *Sales Forecasting Using Neural Networks*, devido a semelhança entre o ambiente encontrado nesses artigos com o estudo realizado por esta dissertação

3.8.2.1. O Problema de Previsão de Venda

No primeiro, *A hybrid econometric-neural network modeling approach for sales forecasting*, os autores relatam a realidade das empresas em torno de suas previsões de venda, afirmando que devido à dinâmica e complexidade do ambiente das corporações, tanto

externos como internos, as empresas têm optado por utilizar modelos qualitativos, em que os resultados costumam ser modelos imprecisos com distúrbios dramáticos em seus planos de produção. Essa é a realidade da empresa onde os autores realizaram a pesquisa (LUXHOJ, RIIS e STENSBALLE, 1996).

O artigo trata sobre a previsão de vendas utilizando RNAs em uma empresa da Dinamarca que produz produtos de áudio e vídeo, apresentando o desenvolvimento de um modelo híbrido, integrando características estruturais dos modelos econométricos com o poder de reconhecimento de padrões não lineares das Redes Neurais Artificiais.

Já, no artigo *Sales forecasting using neural networks*, os autores têm o objetivo de apresentar um método utilizando RNAs para previsão de vendas em uma rede de supermercados, comparando o método proposto com os métodos já utilizados pela empresa.

Conforme os autores deste trabalho, os métodos estatísticos apresentados, até o momento, demonstram certas deficiências, prejudicando uma boa previsão de vendas. Dentre as desvantagens encontradas nesses métodos estão: I) para cada problema, um modelo individual estatístico tem de ser escolhido para fazer alguma suposição sobre as tendências; II) esse tipo de método não tem o poder de prever com precisão modelos com dependências não lineares (THIESING E VORNBERGER, 1997).

Na tentativa de obter soluções alternativas ao problema, os autores aplicaram RNAs na busca pela previsão de vendas de 20 produtos da rede de supermercado, visando sanar os problemas apresentados quanto a utilização de métodos estatísticos com esse objetivo.

Na seção a seguir serão apresentadas as estruturas das RNAs modeladas para cada um dos problemas.

3.8.2.2. Estrutura das RNAs

Em *A hybrid econometric-neural network modeling approach for sales forecasting* a estrutura proposta pelos autores é a de utilizar um modelo híbrido utilizando as variáveis entrada e saída de um modelo regressivo econométrico e outras variáveis em conjunto com uma rede neural artificial com o intuito de capturar comportamento não linear.

A estrutura da RNA utilizada para esse primeiro artigo foi a *Multi-layer Perceptron* com 3 camadas, sendo 11 entradas que são compostas pelas 9 variáveis de entrada do modelo econométrico (mês de previsão, total de vendas no ano anterior, vendas no mês anterior na

Dinamarca, vendas no mês anterior na Noruega, vendas no ano anterior na Suécia, vendas no mês anterior na Suíça, vendas no ano anterior na Itália e previsão exponencial suavizada), previsão de venda do modelo de regressão econométrico e o total vendido pela empresa no mês anterior, 51 neurônios na camada escondida e 1 neurônio na camada de saída.

No artigo *Sales forecasting using neural networks* foi utilizado o algoritmo *Backpropagation* também em uma rede *Multilayer Perceptron*. A estrutura da rede escolhida possui três camadas: uma camada de entrada composta por 7 entradas, sendo duas variáveis contendo a venda de duas semanas antes da data de previsão (uma variável para cada semana), a diferença das vendas da semana atual para a semana anterior ($x_t^f = x_t - x_{t-1}$), um sinalizador de movimento de vendas para a semana, um sinalizador de movimento de vendas para o feriado, indicador de campanhas de marketing e indicador de alterações no preço; uma camada escondida com 4 neurônios e uma camada saída contendo a previsão de venda para a próxima semana.

A estrutura da rede assim como a quantidade de neurônios da camada escondida foram definidas por meio de treinamentos variando o número de neurônios, a taxa de aprendizado, a constante de momento, as variáveis de entrada e o tamanho da janela de tempo.

No item a seguir serão apresentados os resultados obtidos pelos artigos.

3.8.2.3. Resultados

No artigo *A hybrid econometric-neural network modeling approach for sales forecasting* foi realizado uma análise sobre 6 meses, sendo que a aplicação da técnica abordada pelos autores resultou numa modesta redução de 2,3% no erro, se comparado com os métodos qualitativos já utilizados pela empresa. O autor afirma que não foi realizado uma exaustiva procura no espaço de decisão, o que viabiliza a revisão dos modelos utilizados para previsão, possibilitando melhorias nos resultados obtidos.

Em *Sales forecasting using neural networks*, após definida a arquitetura e colhidos os resultados obtidos pelas previsões que utilizaram RNA, foi apresentado um comparativo com os métodos atualmente utilizados pelo supermercado – o método *naive*, que utiliza as vendas do último período como previsão para o período seguinte, e o método *Moving Average*, apresentado no capítulo 2. Nos comparativos realizados, o método utilizando RNA se saiu

melhor que os outros métodos, chegando a uma redução de 17% comparado com o *Moving Average* e 24% quando comparado com o *naive*.

3.8.3. Observações sobre os trabalhos

Na busca de trabalhos relacionados a esta dissertação, não foram encontrados artigos diretamente relacionados a previsão de vendas em indústrias com as mesmas características, por esse motivo, os três artigos analisados, apesar de não serem da mesma área, apresentam métodos para previsão de valores futuros utilizando RNAs, sendo que em *Forecasting of the electric energy demand trend and monthly fluctuation* e *Sales forecasting using neural networks*, o método utilizado como comparativo com as RNAs é o *Moving Average*, mesmo método utilizado atualmente na empresa onde o estudo de caso desta dissertação foi realizada.

O artigo *A hybrid econometric-neural network modeling approach for sales forecasting* foi detalhado com o objetivo de apresentar um caso de insucesso na utilização das RNAs, chegando a uma pequena redução de 2,3% quando comparado aos métodos qualitativos utilizados pela empresa. Mas, conforme relatado pelo autor, o insucesso é dado devido a não utilização de variáveis adequadas que representem a dinâmica de venda de forma correta, assim dificultando a compreensão da dinâmica de venda da empresa pela RNA.

Neste capítulo foram apresentados os conceitos sobre as Redes Neurais Artificiais, fornecendo uma descrição básica da metodologia escolhida para o desenvolvimento desta dissertação. Foram apresentados também trabalhos correlatos, os quais justificam o porquê do desenvolvimento de uma metodologia utilizando RNAs. No capítulo a seguir será detalhada a metodologia desenvolvida para a realização desta pesquisa de mestrado.

4.MODELO ATUAL E EXPERIMENTOS PARA PREVISÃO DE VENDAS

Nesse capítulo serão apresentados os elementos que compõem a experimentação realizada com os dados obtidos na empresa analisada pro esta pesquisa de mestrado. Será apresentada a metodologia aplicada para realização deste trabalho, assim como o modelo quantitativo atualmente utilizado pela empresa para previsão de vendas. Também serão apresentadas a fonte de dados, as ferramentas utilizadas para obtenção dos dados e treinamento da RNA e três experimentos utilizando RNAs, sendo eles: I) Experimento utilizando como dados de entrada para as RNAs os mesmos dados utilizados pelo modelo atual; II) Experimento utilizando outras variáveis, na tentativa de ajudar a rede com dados mais próximos ao dia da previsão; e III) Experimentos utilizando informações referentes à sazonalidade, verificando se esse tipo de informação ajuda a rede a obter um melhor resultado.

4.1.METODOLOGIA

Neste item serão apresentados os elementos que fazem parte da metodologia empregada para realização dos experimentos com objetivo de realizar as previsões de venda.

Na figura 4.1 é apresentada a metodologia na forma de um diagrama:

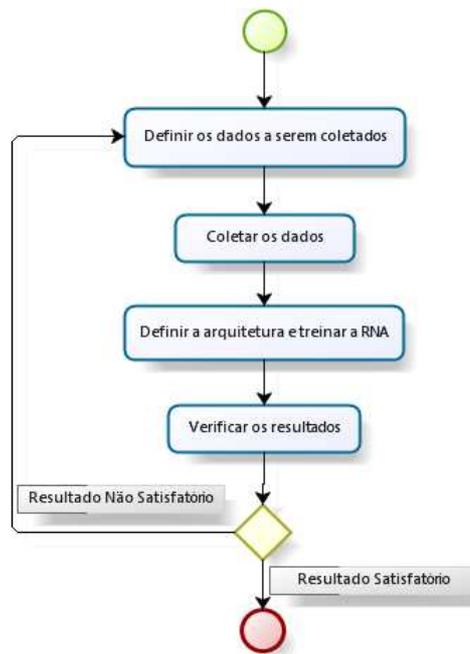


Figura 4.1: Diagrama da Metodologia Proposta

Conforme o diagrama demonstrado na figura 4.1, o primeiro passo da metodologia é a definição dos dados a serem coletados, buscando identificar quais variáveis definem o problema de forma adequada. Após essa etapa, ocorre a coleta de dados no Banco de Dados (BD) da empresa e, na sequência, os dados são utilizados no treinamento da Rede Neural, realizando o treinamento com várias arquiteturas, buscando identificar aquela que mais se ajusta ao problema. Uma vez ocorrido o treinamento, os resultados serão verificados, a fim de se identificar se são satisfatórios, caso não sejam, o processo deverá recomeçar, identificando outras variáveis que poderiam melhorar o desempenho da RNA e então refazendo todas as etapas que vem na sequência. O processo deve ser repetido até que se alcance, se possível, uma RNA que consiga prever as vendas da empresa com um nível aceitável de erro.

No item a seguir é apresentado o modelo estatístico quantitativo atualmente utilizado pela empresa.

4.2.MODELO ATUAL

O modelo quantitativo de Previsão de Vendas utilizado pela empresa em questão tem como base o *Moving Average* (MA), sendo que o MA é rodado três vezes para cada previsão,

com períodos diferentes, no caso o MA dos últimos 12 períodos, MA dos últimos 6 períodos e o MA dos últimos 3 períodos. Cada período é equivalente a um mês.

O modelo utilizado considera somente o maior valor entre os três calculados e utiliza esse resultado como a previsão para o próximo período. Nessa técnica utilizada para previsão de vendas sempre é considerada a previsão mais otimista, ou seja, utiliza o maior valor com o objetivo de que o estoque seja suficiente para que não existam atrasos nos seus pedidos. Entretanto, esse método de considerar o maior valor faz com que a empresa possivelmente tenha um custo maior por estar investindo em uma quantidade de produtos desnecessária para o momento.

Um exemplo para explicar o cenário seria supor a existência dos produtos A, B, C, D e E no dia 01/01/2007, conforme as características apresentadas na tabela abaixo:

Tabela 4.1: Exemplo de informações de produtos utilizadas no Previsão de Vendas

Produto	Média Mensal 12 meses	Média Mensal 6 meses	Média Mensal 3 meses
A	65,50	96,67	171,67
B	240,92	327	300
C	23,75	29,67	29,67
D	23,83	26,67	34,67
E	485,17	547,50	224,67

Considerando o cenário apresentado, a previsão de vendas da empresa realizaria a seguinte previsão para o período posterior a data em que os dados foram coletados:

Tabela 4.2: Resultado da Previsão de Vendas utilizado atualmente na empresa

Produto	Previsão de vendas para o próximo período
A	171,67
B	327,00
C	29,67
D	34,67
E	547,50

Conforme comentado anteriormente, o modelo apresenta problemas em relação a sua previsão. Como o modelo atual utiliza somente a maior média, tende a ter dificuldades em não conseguir prever, por exemplo, que devido às vendas da empresa ocorrerem em sua maior parte dentro do Brasil, no mês de março, por ser o início do ano comercial do país, é um mês

em que ocorre um aumento significativo nas vendas. Contudo, nos meses de janeiro e fevereiro, por ser em meses de férias e de poucas atividades no ramo, ocorre uma baixa significativa nas vendas.

4.3.FONTE DE DADOS

A empresa onde o trabalho foi realizado é classificada como de grande porte no ramo de telecomunicação; é atuante no sul do país e tem 100% dos seus processos mapeados em seu ERP, assim, dispondo de toda história do seu negócio em seu banco de dados – uma base rica de informações a ser explorada.

Quanto ao tema do trabalho, isto é, previsão de vendas de produtos, a empresa utiliza esse tipo de metodologia de dinâmica de previsão de vendas quantitativa desde 2005.

O BD utilizado pela empresa é o Oracle na versão 10g; base de dados consolidada no mercado como uma das principais para armazenamento e gerenciamento de informações em larga escala.

Com o objetivo de consolidar uma base de informações mais sólida sobre a dinâmica de vendas, optou-se pelo uso de dados relativos ao período entre 2006 e 2008 como fonte de informações, pois acredita-se que nesse período de tempo se obtenha um grau mais elevado de confiabilidade das informações para extração devido ao processo de previsão de vendas já estar estável.

4.4.FERRAMENTAS

A seguir serão apresentadas as principais tecnologias e ferramentas utilizadas para obtenção dos dados e treinamento das RNAs:

- Linguagem PL/SQL – é uma linguagem desenvolvida pela Oracle como uma extensão ao SQL (*Structured Query Language*) e tem como objetivo dar mais poder ao desenvolvimento na camada de banco de dados. A linguagem fornece suporte a estruturas básicas como variáveis, condições, vetores, recursos como

Orientação a Objetos (OO), além de uma gama de funções que possibilitam o desenvolvimento de aplicações complexas, seja no banco ou em uma das ferramentas da Oracle que dão suporte a essa linguagem. A Linguagem PL/SQL foi utilizada na etapa de coleta dos dados (ORACLE CORPORATION, 2010).

- PL/SQL Developer – ferramenta desenvolvida pela All Around Automations que busca dar agilidade ao desenvolvimento utilizando PL/SQL, trazendo inúmeras facilidades ao desenvolver na linguagem. O principal foco da ferramenta é a facilidade no uso, a qualidade do código e a produtividade, palavras-chaves quando se trata do desenvolvimento de aplicações em Oracle PL/SQL. A ferramenta PL/SQL Developer também foi utilizada na etapa de coleta dos dados (ALLROUND AUTOMATIONS, 2010).
- R-Project – é uma linguagem e um ambiente para computação estatística e gráfica. O software é de código livre e possibilita, entre seus diversos recursos, o treinamento de RNAs, análise estatística e gráfica. Uma das principais forças do R é facilidade na geração de gráficos, incluindo símbolos matemáticos e fórmulas, quando necessário. O R-Project foi utilizado para normalizar os dados a serem utilizados na RNA e para o treinamento da RNA (R DEVELOPMENT CORE TEAM, 2009).
- AMORE (*A More Flexible neural network package*) – é um pacote desenvolvido para o R-Project com o objetivo de dar mais flexibilidade aos usuários da ferramenta quanto a utilização de RNAs e acesso aos seus parâmetros, provendo mais controle no treinamento. Essa ferramenta também foi utilizada na etapa de definir a arquitetura e treinar a RNA (LIMA, 2010).

4.5.COLETA DOS DADOS

Através da ferramenta PL/SQL Developer, foi coletado diretamente do BD da empresa estudada os dados de vendas referentes aos anos de 2006, 2007 e 2008, em que, para cada produto, foi obtida a média de venda mensal dos 12 meses, dos 6 meses e dos 3 meses antecedentes a data da Previsão de Venda. Foram coletados também a quantidade vendida dos produtos em períodos fechados próximos ao problema, sendo eles 90 a 61 dias, 60 a 31 dias e

30 a 1 dias que antecedem a previsão e também a média mensal de produtos vendidos considerando o período de um mês após a data da previsão.

Abaixo, um exemplo dos dados coletados de cinco diferentes produtos em uma Previsão de Venda realizada em 01/01/2007:

Tabela 4.3: Exemplo de dados utilizados pelo *Forecasting*

ID	Média Mensal 12 meses	Média Mensal 6 meses	Média Mensal 3 meses	90 a 61 dias antes previsão	60 a 31 dias antes previsão	31 a 1 dias antes previsão	Venda 30 dias após data previsão	Data <i>Forecasting</i>
A	31,5	47,16	51,33	68	54	32	48	01/01/2007
B	176,75	233	286	270	310	278	240	01/01/2007
C	6,91	10,66	8,66	8	11	7	9,78	01/01/2007
D	238,33	324,83	415,33	400	422	424	267,39	01/01/2007
E	58,33	116,66	168,66	180	172	154	38,91	01/01/2007

Foi coletado um total de 2.461 registros, estando eles entre os meses de janeiro de 2006 a dezembro de 2008. Para o treinamento não foram utilizados os IDs (identificador do produto) como entrada da rede. Após uma análise sobre os dados, entendeu-se que questões sazonais e outros aspectos que poderiam variar entre os produtos são similares para toda a linha de produtos da empresa, assim conclui-se que somente com as informações sobre as vendas, a RNA pode chegar a uma compreensão generalizada da dinâmica de venda de qualquer produto da empresa. As informações do mês em que a Previsão de Venda estava sendo executada foi utilizada em um dos experimentos realizados, a fim de compreender se essa informação agregava valor ao cenário apresentado. Nos dados coletados não constavam informações qualitativas, que poderiam ajudar no treinamento da RNA. Por exemplo, caso a empresa assine um novo contrato com um grande cliente, a demanda aumentaria em uma quantia muito elevada, portanto alterando a dinâmica de vendas de forma a não possibilitar a predição com a RNA. O oposto também traria o mesmo problema, ou seja, se empresa perdesse um grande cliente, a demanda de seus produtos teria uma queda muito grande, também impossibilitando a rede de conseguir prever a demanda.

4.6. TÉCNICAS PARA AVALIAÇÃO DOS RESULTADOS

Neste item serão apresentadas as técnicas utilizadas para avaliar a qualidade das previsões, por meio delas foi possível identificar qual a melhor arquitetura para resolução do problema em cada um dos experimentos.

Existem diversas técnicas que buscam identificar a qualidade da rede, como por exemplo, o *holdout*, o *jackknife* e o *k-fold cross validation*. O método *k-fold cross validation*, que foi utilizado nesta dissertação, será apresentado na seção a seguir.

4.6.1. K-fold Cross Validation

K-fold cross validation é um dos métodos mais comuns para classificação de precisão, ele permite apurar quão preciso está sendo a previsão de um classificador.

Segundo Han e Kamber (2005), no *k-fold cross validation* os dados são randomicamente particionados em k partes, S_1, S_2, \dots, S_k , aproximadamente de mesmo tamanho. Treinamento e testes são realizados k vezes. Na interação i , a parte S_i é reservada para teste e o restante das partes são coletivamente utilizadas para treinar o classificador. Ou seja, o classificador da primeira interação é treinado com as partes S_2, \dots, S_k , e testado com S_1 ; o classificador da seção treinada com as partes S_1, S_3, \dots, S_k é testada com S_2 ; e assim por diante.

Em geral, o *10-fold cross validation* é o método mais utilizado para estimar a precisão de um classificador, dividindo os dados em 10 partes, utilizando 9 para treinamento e 1 para teste em cada rodada.

Ao final da execução toda a população foi utilizada para validar a qualidade do classificador. O método escolhido para averiguar o erro obtido em cada parte do teste é o MAPE, que será apresentado no item a seguir.

4.6.2.MAPE

O *Mean Absolute Percentage Error* (MAPE) é o método mais utilizado para medir a precisão de uma previsão. O MAPE expressa o percentual de erro entre a previsão e o valor real para o período a ser testado e tem o benefício de ser relativamente fácil de se compreender e ser correlacionado a resultados do negócio (GATTORNA, 1998).

O método MAPE pode ser definido por:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right|$$

Onde:

F_t = Previsão para o período t ,

A_t = Valor real para o período.

No caso do trabalho realizado, o MAPE foi utilizado em conjunto com o *10-fold cross validation* para avaliar a precisão do resultado a cada rodada.

Nos próximos itens serão apresentados os experimentos utilizando RNAs, em que a RNA a partir dos dados fornecidos, tentará compreender a dinâmica das vendas e então realizar a previsão para o período seguinte.

4.7.EXPERIMENTO 1

O primeiro experimento consiste em utilizar como base os dados utilizados no modelo quantitativo da empresa, isto é, será um modelo híbrido, que ao invés de utilizar o maior valor entre MA 12, MA 6 e MA 3, esses valores foram utilizados como entrada para a RNA, e como saída, será utilizado a quantidade vendida real no período posterior a data de obtenção dos dados.

Um exemplo dos dados utilizados para o treinamento da RNA é apresentado na tabela abaixo:

Tabela 4.4: Exemplo de dados utilizados no Experimento 1

Data da Previsão de Vendas	MA 12 meses	MA 6 meses	MA 3 meses	Vendas 01/2007
01/01/2007	18	61	19	10
01/01/2007	98	52	17	32
01/01/2007	377	132	175	204
01/01/2007	34	23	20	24
01/01/2007	219	377	362	241

Os dados utilizados no treinamento foram normalizados entre 0 e 1 por ser uma exigência da função sigmóide, que foi utilizada como função de ativação. A definição de utilização da função sigmóide foi empírica.

Esse primeiro experimento tem como principal objetivo realizar um comparativo entre o método convencional utilizando o maior valor entre os três experimentos empregando MA e uma RNA treinada utilizando essas mesmas entradas.

A seguir será exposto o processo de definição da arquitetura da RNA para o Experimento 1 e o treinamento realizado.

4.7.1. Modelagem e Treinamento da RNA

Buscando definir a melhor arquitetura da RNA para a resolução do problema, realizaram-se diversos treinamentos utilizando o *10-Fold Cross Validation*. Para cada treinamento foi utilizado uma quantidade diferente de neurônios na camada escondida, variando entre 1 a 15 neurônios. O método para validação de qual a arquitetura com menor erro, conseqüentemente com melhor desempenho, foi o MAPE, sempre validando com o *Fold* que ficou fora do treinamento a cada rodada.

Uma vez realizados todos os treinamentos e avaliados os resultados obtidos, foi encontrado como melhor arquitetura para o problema um modelo utilizando 3 neurônios na camada de entrada (MA 12, MA 6 e MA 3), 13 neurônios na camada escondida, os quais foram definidos pelo método citado acima, e um neurônio na camada de saída para o período seguinte.

Abaixo é mostrada uma representação gráfica da arquitetura da RNA modelada para o problema:

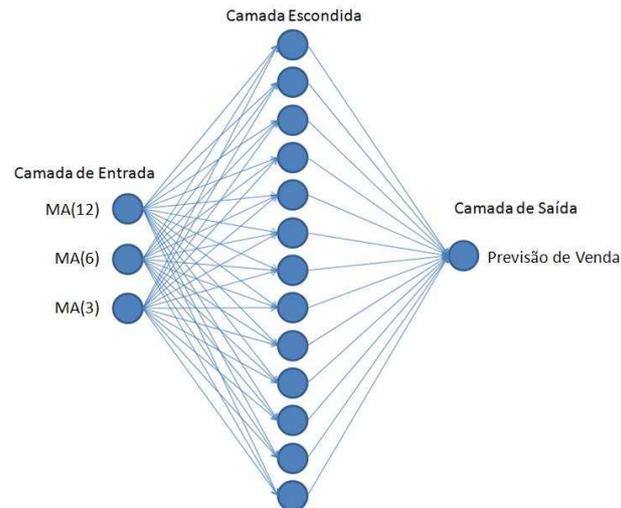


Figura 4.2: Arquitetura da RNA do Experimento 1

No capítulo 5 serão apresentados os resultados e um comparativo entre o Experimento 1 e a adaptação do método MA utilizado pela empresa. No anexo 1 são apresentados exemplos dos dados utilizados para validar a RNA treinada no Experimento 1, assim como os resultados obtidos.

4.8.EXPERIMENTO 2

Para o Experimento 2, utilizou-se dados mais próximos ao período da realização da previsão, entendendo que o período de 12 meses é muito distante do problema, por exemplo, nos primeiros 8 meses a empresa poderia ter um contrato em que fazia as vendas de determinado produto serem elevadas, mas nos últimos quatro meses, caso o contrato tenha sido encerrado, a venda desse produto será muito mais baixa, o que pode ocasionar distúrbios no momento da previsão utilizando a média dos 12 meses, trazendo mais ruído para a rede do que benefício, apresentando dados de uma realidade passada que teve uma grande alteração no período recente. Outro problema propenso a ocorrer é a utilização de médias, trazendo uma carga de informação de um período para outro, por exemplo, a empresa pode ter vendido 120 unidades no mês anterior e somente 4 no mês atual, se for feita a média, conforme os métodos até agora apresentados, ficaria uma média de 62 unidades vendidas por mês, o que não representa a realidade da dinâmica desse produto, que sofreu uma forte baixa no período atual.

Buscando uma alternativa aos problemas apresentados acima, foram considerados somente os últimos três meses como variáveis de entrada, a fim de se obter informações próximas a data de previsão, tratando cada um deles como períodos isolados, sem utilizar médias, ou seja, as variáveis serão do período entre 90 e 61 antes da previsão, 60 e 31 dias antes da previsão e 30 e 1 dias antes da previsão, com isso obtendo três variáveis com cada uma contendo a carga de vendas de períodos isolados e próximos a data de previsão.

Na tabela abaixo serão exemplificadas as informações utilizadas no Experimento 2:

Tabela 4.5: Exemplo de dados utilizados no Experimento 2

Data da Previsão de Vendas	Vendas 10/2006 (3 m antes)	Vendas 11/2006 (2 m antes)	Vendas 12/2006 (1 m antes)	Vendas 01/2007
01/01/2007	18	61	19	10
01/01/2007	98	52	17	32
01/01/2007	377	132	175	204
01/01/2007	34	23	20	24
01/01/2007	219	377	362	241

No caso da tabela 4.5, as vendas dos meses 10, 11 e 12 de 2006 são as informações utilizadas como entrada para o treinamento, e como saída é utilizado o período de 01/2007. Assim como no Experimento 1, os dados utilizados também foram normalizados entre 0 e 1 para a realização do treinamento.

No item a seguir serão apresentados a arquitetura encontrada e o treinamento realizado no Experimento 2.

4.8.1. Modelagem e Treinamento da RNA

Assim como no Experimento 1, no Experimento 2 foi utilizado também o *10-Fold Cross Validation* para definir a melhor arquitetura, tendo o MAPE como método de avaliação dos resultados.

A melhor arquitetura da RNA encontrada para o Experimento 2 foi idêntica a do Experimento 1, isto é, 3 neurônios na camada de entrada, sendo, dessa vez, as unidades vendidas nos períodos de 90 a 61 dias, 60 a 31 dias e 30 a 1 dias que antecedem o dia da

previsão, 13 neurônios na camada escondida, e um neurônio na camada de saída o qual representa a venda para o período desejado.

Abaixo é apresentado graficamente a arquitetura do Experimento 2:

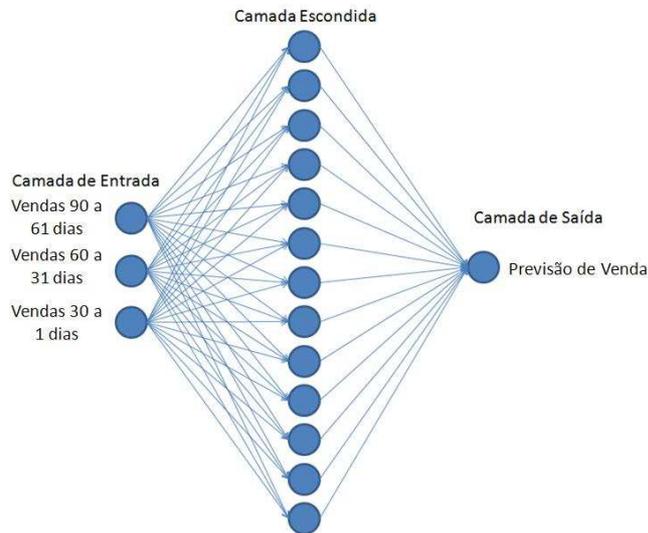


Figura 4.3: Arquitetura da RNA do Experimento 2

No capítulo 5 será apresentado um comparativo entre o Experimento 1 e o Experimento 2, buscando identificar se houve melhoria quanto a mudança das variáveis utilizadas pelo problema. No anexo 2 são apresentados exemplos dos dados utilizados para validar a RNA treinada no Experimento 2, assim como os resultados obtidos.

4.9.EXPERIMENTO 3

No Experimento 3 foi utilizada a mesma estrutura do Experimento 2, mas adicionando como informação para a RNA o mês em que a previsão está sendo realizada. O objetivo, com isso, é ajudar a rede a compreender questões sazonais, sabendo que se uma previsão for feita, por exemplo, em janeiro, a previsão deverá ser mais baixa que, por exemplo, uma previsão no mês de julho.

Foi utilizada uma variável para cada mês, podendo o valor ser igual a 1, o que significa que esse mês é o mês corrente, ou igual a 0, não estando no mês da variável. Apenas um mês ficará com valor igual a 1 por linha.

Abaixo é ilustrado um exemplo das informações utilizadas pela RNA incluindo informações sazonais:

Tabela 4.6: Exemplo de dados utilizados no Experimento 3

Data da Previsão de Vendas	01	02	03	04	05	06	07	08	09	10	11	12	Vendas 1 mês antes	Vendas 2 m antes	Vendas 3 m antes	Vendas Período
01/01/2007	1	0	0	0	0	0	0	0	0	0	0	0	377	132	175	204
01/08/2007	0	0	0	0	0	0	0	1	0	0	0	0	123	152	134	103
01/04/2007	0	0	0	1	0	0	0	0	0	0	0	0	78	61	43	86
01/06/2007	0	0	0	0	0	1	0	0	0	0	0	0	51	89	61	58
01/12/2007	0	0	0	0	0	0	0	0	0	0	0	1	38	40	35	15

No caso do exemplo acima, as variáveis de quantidade vendida e as de vendas no período foram normalizadas entre 0 e 1, assim como no Experimento 1 e 2. As variáveis referentes ao mês em que a Previsão de Vendas foi realizada não necessitaram de ser normalizadas por já estarem dentro dos valores possíveis de serem trabalhados com a função sigmóide.

No item a seguir será apresentada a arquitetura encontrada pela rede e o treinamento realizado.

4.9.1. Modelagem e Treinamento da RNA

Assim como nos experimentos anteriores, foi utilizado o *10-Fold Cross Validation* em combinação com o MAPE buscando identificar a melhor arquitetura para resolver o problema.

A melhor arquitetura encontrada foi por meio da utilização de 15 entradas, sendo 12 delas as informações referentes ao mês e as outras 3 referentes às unidades vendidas nos períodos de 90 a 61 dias, 60 a 31 dias e 30 a 1 dias que antecedem a previsão, 15 neurônios na camada escondida e um neurônio na camada de saída.

A representação gráfica da arquitetura da RNA modelada para o problema é apresentada abaixo:

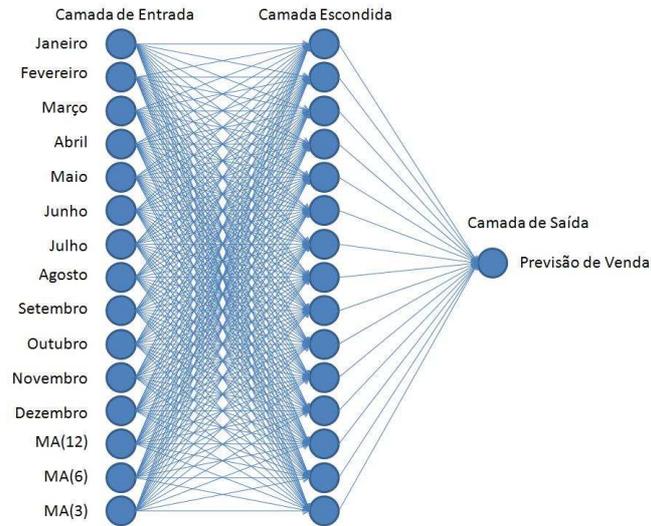


Figura 4.4: Arquitetura da RNA do Experimento 3

Os resultados dos 3 experimentos realizados, além de um comparativo entre eles, a fim de identificar o desempenho das RNAs no problema apresentado e reconhecer se as otimizações trouxeram melhorias às previsões realizadas, serão demonstradas no capítulo 5. No anexo 3 são apresentados exemplos dos dados utilizados para o validar a RNA treinada no Experimento 3, assim como os resultados obtidos.

5.RESULTADOS

Nesse capítulo serão apresentados os resultados obtidos em cada experimento assim como um comparativo entre os métodos, buscando identificar se as RNAs são alternativas ao problema tratado nesta dissertação, assim como avaliar se a utilização de outras variáveis ajudariam a RNA a compreender melhor a dinâmica de venda da empresa.

5.1.EXPERIMENTO 1

Após identificada a arquitetura ideal para a solução do problema, foram analisados os resultados obtidos no Experimento 1, obtendo-se um erro médio de 317% através da análise utilizando o *10-fold Cross Validation* em conjunto com o MAPE. O menor erro encontrado foi de 199% e o maior erro foi de 693%. A figura abaixo apresenta o erro obtido para cada *fold*.

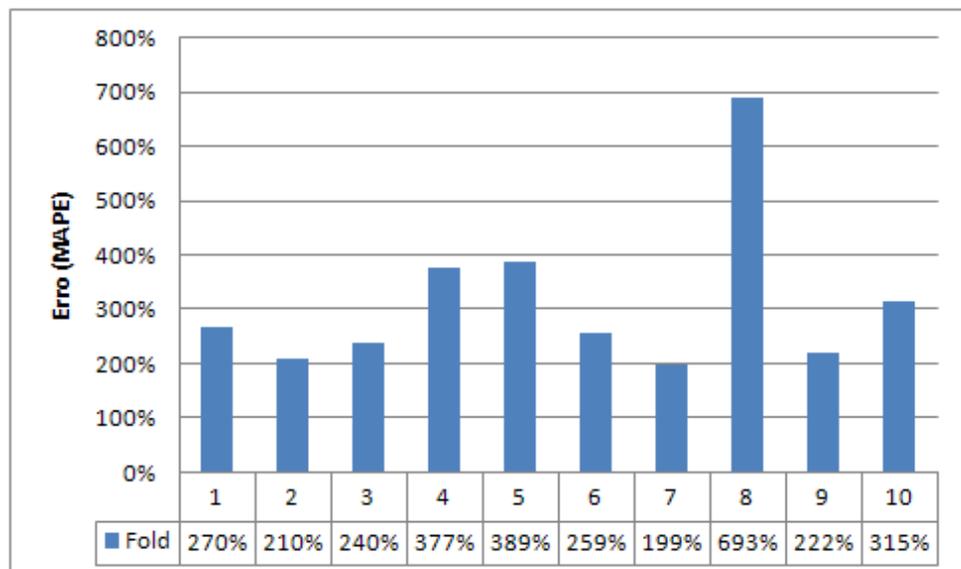


Figura 5.1: Erro da RNA por *fold* obtido no Experimento 1.

Realizando uma análise somente baseada no experimento, o resultado não mostrou-se satisfatório para resolver o problema de previsão de venda, uma vez que, com os percentuais apresentados, o erro nas previsões permaneceria muito elevado. Como medida para avaliar a

real eficácia do método utilizando RNA, é apresentado abaixo os resultados obtidos pelo método utilizado pela empresa (de onde foram retiradas as variáveis utilizadas no treinamento da RNA para o Experimento 1, assim permitindo um comparativo claro sobre o desempenho de cada um dos métodos). Os resultados foram coletados a cada rodada do método *10-fold Cross Validation*, possibilitando um comparativo *fold* por *fold*. O MAPE, utilizado para avaliar o erro da RNA, também foi empregado para avaliar o erro do método de que se serviu a empresa, permitindo uma análise com o mesmo sistema de medida.

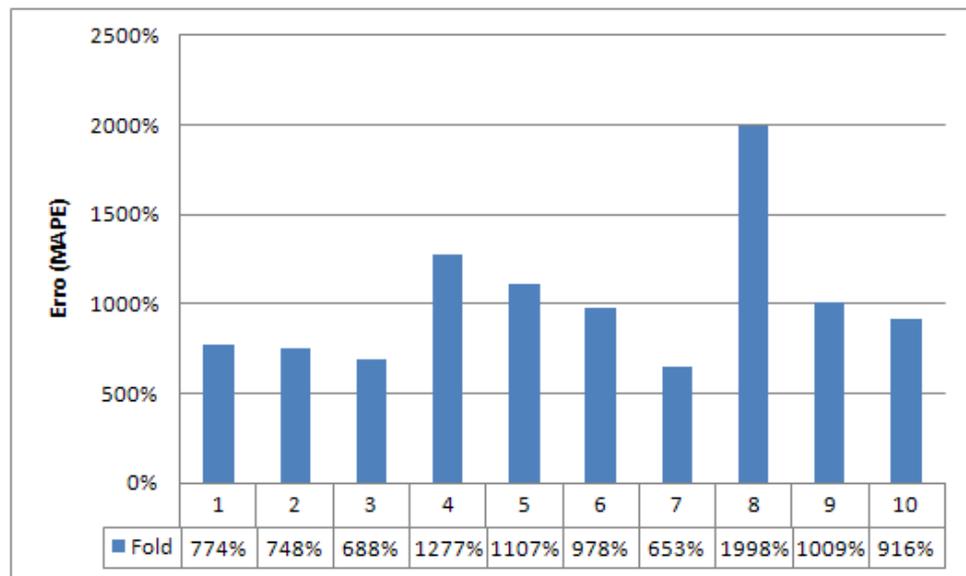


Figura 5.2: Erro do método tradicional por *fold* obtido no Experimento 1.

Conforme apresentado na figura acima, pode-se identificar que o erro do método tradicional é muito superior àquele obtido com a RNA, chegando a um percentual máximo de 1998% e mínimo de 653%, e um erro médio de 1015%.

Realizando uma análise sobre os gráficos apresentados, é facilmente identificável que o *fold* 8 aquele em que os dois métodos tiveram o pior resultado. Através dessa análise pode-se compreender que a informação contida nesse *fold* não representa claramente as vendas realizadas, justificando o alto erro da RNA especificamente nesse caso ou identificando a ausência de informação que ajudasse a RNA a compreender a diferente dinâmica ocorrida.

Comparando os dois métodos, em todos os *folds* a RNA se saiu melhor que o método atualmente utilizado pela empresa, chegando a uma redução do erro médio em torno de 698%. Apesar do resultado com a RNA ainda não ser o desejado, visando o problema da empresa quanto as suas previsões de venda, em relação ao método atualmente utilizado, os ganhos são

consideráveis, demonstrando que as variáveis utilizadas, extraídas do método tradicional, não são suficientes ou adequadas para o problema.

5.2.EXPERIMENTO 2

Após traçadas as novas variáveis e definida a nova arquitetura da RNA, foram coletadas as respostas da rede a cada *fold*, buscando identificar se os resultados chegaram a um nível satisfatório. Para o Experimento 2 o erro médio obtido ficou em 265%, sendo o menor percentual de 176% e o maior de 512%. Este, assim como no Experimento 1, também está localizado no *fold* 8. Os resultados do Experimento 2 para cada *fold* são apresentados na figura 5.3

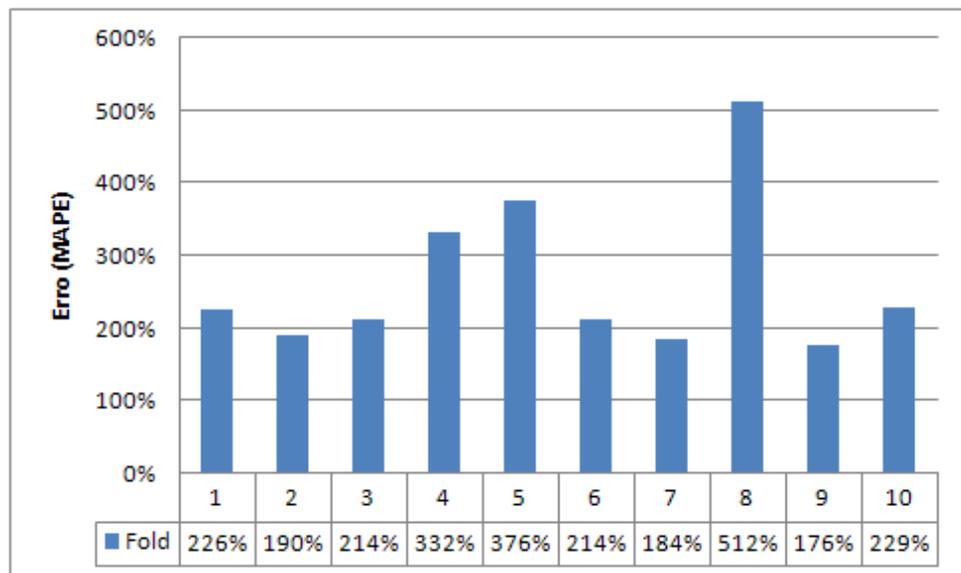


Figura 5.3: Erro da RNA por *fold* obtido no Experimento 2.

Comparando os resultados do Experimento 1 com os resultados do Experimento 2, é possível avaliar que as informações de venda mais próximas a data da previsão representam melhor as vendas para o próximo período, conseguindo reduzir o erro em comparação ao Experimento 1, cerca de 52%, e reduzir o erro em relação ao método atualmente utilizado pela empresa em torno de 750%.

Os resultados obtidos no Experimento 2, apesar de alcançarem resultados muito superiores ao método tradicional, em relação ao problema, ainda representa um erro muito

elevado para quem necessita realizar um planejamento em relação as suas vendas, não chegando ainda a um nível desejado pela empresa.

5.3.EXPERIMENTO 3

No Experimento 3, as entradas são as mesmas do Experimento 2, mas com a adição da sazonalidade, buscando ajudar a RNA a compreender a dinâmica específica do período em que a previsão está sendo realizada.

Os resultados foram sutilmente superiores àqueles obtidos no Experimento 2, o erro médio ficou em 261%, o máximo em 495% e o mínimo em 175%. Em diversos *fold*s o erro foi superior ao obtido no Experimento 2, mostrando que aparentemente as informações do período em que a previsão está sendo realizada não têm uma contribuição significativa no resultado geral, pelo menos com a estrutura de dados utilizada. Na figura 5.4 são apresentados os resultados obtidos em cada *fold* no Experimento 3.

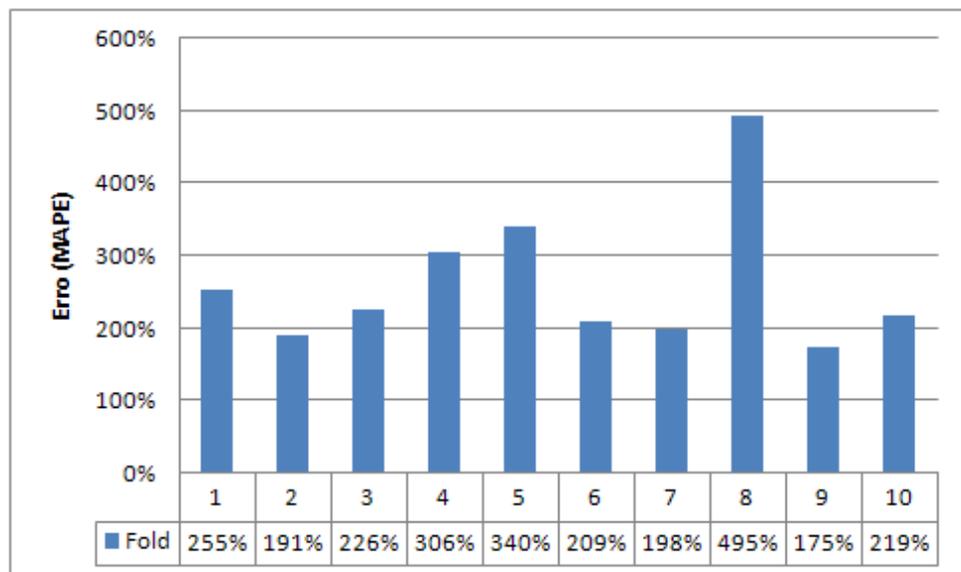


Figura 5.4: Erro da RNA por *fold* obtido no Experimento 3.

Avaliando o resultado do Experimento 3, que apesar de não alcançar uma melhoria significativa, se comparado ao Experimento 2, ainda assim foi o melhor resultado alcançado entre os 3 experimentos, atingindo a uma redução do erro em 754% comparado ao método utilizado pela empresa.

Neste capítulo foram apresentados os resultados alcançados nos três experimentos realizados com os dados obtidos da empresa, realizando um comparativo com o método atualmente utilizado pela empresa. Por fim, as conclusões obtidas com a realização da pesquisa e as expectativas para trabalhos futuros serão apresentadas, enfim.

6.CONCLUSÕES

A partir da necessidade de uma indústria do ramo de telecomunicação em prever as suas vendas, esta dissertação teve a oportunidade de realizar um estudo sobre um ambiente rico de informações, com processos 100% informatizados, mas, ao mesmo tempo, carente de ferramentas capazes de auxiliarem os gestores no momento da tomada de decisão.

Uma vez que o principal objetivo das indústrias de eletrônicos em geral é a produtividade, identifica-se um grande investimento tecnológico e em TI, no controle de processos de fabricação de produtos. Entretanto, existe uma carência por ferramentas que auxiliem num planejamento adequado de vendas, que é o fator que fornece os melhores indicativos de diagnóstico de empresas produtivas. O desenvolvimento de ferramentas eficientes que ofereçam suporte à tomada de decisão é uma necessidade constante de empresas com este perfil.

Considerando o foco deste trabalho, nas previsões de vendas realizadas na indústria estudada, assim como nas indústrias em geral, são utilizados métodos muito simplificados que não oferecem níveis de confiança significativos aos tomadores decisão, fazendo com que eles utilizem com muita frequência métodos qualitativos, que trazem consigo um nível de precisão muito abaixo do desejado.

Na busca por trabalhos relacionados a área tema, isto é, indústrias que produzam equipamentos de telecomunicação ou eletrônicos com ambiente parecido com o estudado, praticamente não foi encontrado nada relacionado, abrindo nessa área uma oportunidade de realização de trabalhos que agreguem real valor ao ambiente.

Devido à vontade e à necessidade da empresa em melhorar seu processo de previsão de vendas a interação foi muito agradável, visto que a empresa disponibilizou todos os recursos necessários para o trabalho, sejam eles dados, informações, experiência ou mesmo conhecimento, tornando a realização da pesquisa muito mais produtiva.

Comparando a metodologia proposta por esta dissertação com o método tradicional, observou-se que naquela as RNAs se mostraram muito satisfatórias, chegando a reduções significativas do erro. Mas pensando na realidade da empresa, os resultados ainda trazem um grau de incerteza muito elevado quanto as suas vendas futuras. O método quantitativo que a empresa utiliza é bastante simples, apenas suavizando os períodos anteriores através do cálculo da média do período. Já existem evoluções desse método, como o método *Auto*

Regressive Moving Average (ARIMA), o qual vem sendo amplamente utilizado (ZANG, 2003). A simplicidade e a imprecisão do método adotado pela empresa para realizar as previsões de venda, serviram de incentivo para a utilização de um método mais robusto e de comprovado sucesso na literatura, no caso as RNAs para obtenção de resultados significativos, apresentando uma redução do erro em torno de 754%, mostrando ser muito superior inclusive no experimento que utilizou as mesmas variáveis do método atual.

O modelo atualmente aplicado, que serviu de base de informação para o treinamento das RNAs, utiliza somente as vendas passadas como informação para prever as vendas futuras. Assim, o Experimento 1 emprega essas mesmas variáveis para o treinamento. Já nos experimentos 2 e 3 o objetivo foi utilizar a mesma estrutura, ou seja, somente histórico de vendas, apenas otimizando essas variáveis, no caso do Experimento 2, trazendo informações mais próximas a data de previsão e no Experimento 3, trazendo informações sazonais. Pelos resultados obtidos, apesar de muito superiores ao método simplificado utilizado pela empresa, observa-se claramente que as variáveis utilizadas pelo problema não são suficientes para representar a dinâmica de vendas da empresa, necessitando, em primeiro lugar, de um estudo sobre o mercado em que a empresa está inserida e sobre quais as variáveis representam de forma adequada a dinâmica deste.

O trabalho e a pesquisa realizada abrem espaço para estudos que busquem se aprofundar sobre a questão de previsão de vendas nas indústrias relacionadas ao assunto, desde a compreensão do cenário – foco do trabalho –, analisando variáveis internas e externas que influenciam as vendas nesse setor, até a análise dos métodos existentes capazes de realizar previsões com um nível de erro baixo, que permitam às empresas organizarem de forma mais adequada seus recursos materiais e humanos, aumentando sua liquidez e reduzindo suas incertezas.

Daqui em diante, o objetivo é dar seguimento à pesquisa realizando novos experimentos, analisando um possível aumento do volume de dados utilizado para treinamento da RNA, analisando outras técnicas que podem otimizar os resultados, assim como avaliando outras variáveis contidas na base de dados da empresa que porventura podem contribuir para melhoria dos resultados, visando atingir o nível de precisão desejado pela empresa. Esta enseja utilizar um sistema dessa natureza em suas previsões de venda; portanto, uma vez que os resultados cheguem a um nível aceitável, visa a sua implantação em curto prazo.

REFERÊNCIAS

AFKHAMI, R.; YAZDI, F. M. **Application of Neural Networks for Short-Term Load Forecasting**. Power India Conf., 2006.

Allround Automations. **PL/SQL Developer**. Disponível em www.allroundautomations.com último acesso em: jan. 2010.

ALLON, I.; QI, M.; SADOWSKI, R. J. **Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods**. Journal of Retailing and Consumer Services, v. 8, 147--156, 2001.

ARDIL, C. **A Multi-Layer Artificial Neural Network Architecture Design for Load Forecasting in Power Systems**. International Journal of Applied Mathematics and Computer Sciences, v. 4, n. 4, 2007.

CAVALHEIRO, D. **Método de previsão de Demanda aplicada ao Planejamento da Produção de Indústrias de Alimentos**. Dissertação de Mestrado – Programa de Pós graduação em Engenharia de Produção. Porto Alegre: UFRGS, 2003.

CORRÊA, H. L.; GIANESI, I.G.N.; CAON, M. **Planejamento, Programação e Controle de Produção**. 4 ed. São Paulo: Atlas, 2001.

CYBENCO, C. **Approximation by superpositions of a sigmoidal function**. Mathematics of Control, Signals and Systems. v. 2, 303--314, 1989.

DAVIS, M. M.; AQUILANO, N. J.; CHASE, R. B.; **Fundamentos da Administração da Produção**. 3 ed. Porto Alegre: Bookman, 2001.

DOGANIS, P.; ALEXANDRIDIS, A.; PATRINOS, P.; SARIMVEIS, H. **Time series sales forecasting for short shelf-life food products based on artificial neural networks and evolutionary computing**. Journal of Food Engineering. v. 75, 196--204, 2006.

GATTORNA, J. L. **Strategic supply chain alignment: best practice in supply chain management**. 6 ed. Burlington: Gower Publishing Company, 1998.

GONZÁLES-ROMERA, E.; JARAMILLO-MORÁN, M.A.; CARMONA-FERNÁNDEZ, D. **Forecasting of the electric energy demand trend and monthly fluctuation with neural networks**. Computer & Industrial Engineering. v. 52, 336--343, 2007.

HAN, J.; KAMBER, M.: **Data Mining Concepts and Techniques**. 2 ed. San Francisco: Morgan Kaufmann, 2005.

HAYKIN, S. **Redes Neurais: Princípios e prática**. 2 ed. Porto Alegre: Bookman, 2001.

HEBB, D. **The Organization of Behavior: A neurophysiological theory**. New York: Willey, 1949.

KOTLER, P. Administração de Marketing. **Análise, Planejamento, Implementação e Controle**. 5 ed. São Paulo: Atlas, 1998.

KOVÁCS, Z. L. **Redes Neurais Artificiais: Fundamento e Aplicações**. 2 ed. São Paulo: Edição Acadêmica, 1996.

KRAJEWSKI, L. J.; RITZMAN, L. P. **Operations Management, Strategy and Analysis**. 5 ed. MA: Addison-Wesley, Reading, 1999.

KUO, R. J. **A sales forecasting system based on fuzzy neural network with initial weights generated by genetic algorithm**. European Journal of Operational Research, v. 129, 496--517, 2001.

LAUDON, K. C.; LAUDON, J. P. **Sistemas de informação com internet**. Rio de Janeiro: JC Editora, 1999.

LIMA, M. C. **A MORE flexible neural network package**. Disponível em wiki.r-project.org/rwiki/doku.php?id=packages:cran:amore último acesso em: jan. 2010.

LUXHOJ, J. T.; RIIS, J. O.; STENSBALLE, B. **A hybrid econometric-neural network modeling approach for sales forecasting**. International Journal of Production Economics. v. 43, 175--192, 1996.

MAKRIDAKIS, S.; WHEELWRIGHT, S.; HYNDMAN, R. J. **Forecasting methods and applications**. 3 ed. New York: John Wiley & Sons, 1998.

MENTZER, J. T.; COX Jr., J. E. **Familiarity, application, and performance of sales forecasting techniques**. Journal of Forecasting, v. 3, 27--37, 1997.

MOREIRA, D. A. **Administração da Produção e Operações**. São Paulo: Pioneira, 1996.

NAHMIAS, S. **Production and Operation Analysis**. Ontario: Irwin, 1993.

Oracle Corporation. Oracle PL/SQL 11G. Disponível em www.oracle.com/technology/tech/pl_sql/index.html último acesso em: jan. 2010.

ORTIZ-ARROYO, D. **Accurate Electricity Load Forecasting with Artificial Neural Networks**. Computational Intelligence for Modelling, Control and Automation, v.1, 2005.

R Development Core Team. **R-Project is a free software environment for statistical computing and graphics**. Disponível em www.r-project.org último acesso em: jun. 2009.

ROSENBLAT, F. **Principles of Neurodynamics**. Washington D.C: Spartan Books, 1962.

RUSSELL, S.; NORVIG, P. **Artificial Intelligence: a modern approach**. Prentice Hall, 1995.

SHIRVANY, Y.; HAYATI, M. **Artificial Neural Network Approach for Short Term Load Forecasting for Illam Region**. International Journal of Electrical, Computer, and Systems Engineering, v. 1, n. 2, 2007.

STAIR, R. M. **Sistemas de Informação: uma abordagem gerencial**. Rio de Janeiro: LTC, 1998.

THIESING, F. M.; VORNBERGER, O. **Sales Forecasting Using Neural Networks**. ICNN, v. 4, 2125-2128, 1997.

TSEKOURAS, G. J. **An Optimized Adaptive Neural Network for Annual Midterm Energy Forecasting**. Power Systems, IEEE T. on, v.21, n.1, 2006.

TURBAN, E.; McLEAN, E.; WETHERBE, J. **Tecnologia da Informação para Gestão: transformando os negócios na economia digital**. 3. ed. Porto Alegre: Bookman, 2004.

WU, C. H.; MCLARTY, J. W. **Neural Networks and genome informatics**. New York: Elsevier, 2000.

YAFEE, R. **Introduction to Time Series Analysis and Forecasting: with application of SAS an SPSS**. New York: Academic Press, 2000.

ZANG, G. P. **Time series forecasting using hybrid ARIMA and Neural Network model**. Neurocomputing, v. 50, 159--175, 2003.

ANEXO 1 – EXEMPLO DE ENTRADAS E SAÍDAS DO EXPERIMENTO 1

Na tabela abaixo são apresentados exemplos das entradas utilizadas no treinamento do Experimento 1, do valor real vendido, das previsões realizadas pela RNA e do percentual de erro MAPE para cada previsão. Os dados foram alterados por um fator de multiplicação com o objetivo de resguardar a segurança das informações da empresa onde o trabalho foi realizado.

Tabela A: Entradas e saídas do Experimento 1.

MA(12)	MA(6)	MA(3)	Qtd. Vendida	Previsão Exp. 1	MAPE Exp. 1
2,18	4,36	8,73	5,28	5,56	0,05
1607,27	1571,11	1753,44	358,96	486,65	0,36
55,17	36,15	3,12	0,41	5,27	11,97
1543,22	1202,72	1353,26	1336,24	414,13	0,69
19,32	15,90	16,83	1,22	10,19	7,36
225,65	235,62	246,84	2,85	59,91	20,05
19,79	39,58	79,16	24,39	34,50	0,41
25,25	48,93	46,13	15,04	24,76	0,65
11,06	22,13	44,26	53,66	20,30	0,62
20,57	33,97	25,56	47,16	15,43	0,67
4,83	9,66	19,32	20,73	10,04	0,52
83,84	102,54	111,58	16,26	53,99	2,32
99,11	87,58	56,10	14,23	32,02	1,25
74,33	24,00	14,96	11,38	10,36	0,09
47,22	94,44	130,28	163,02	59,49	0,64
153,18	214,43	191,99	133,34	92,94	0,30
32,73	43,01	38,02	15,04	21,07	0,40
48,46	73,24	64,83	27,24	34,21	0,26
11,22	10,29	6,86	39,43	5,75	0,85
12,00	12,78	13,09	12,60	8,43	0,33
410,47	310,11	122,80	15,04	73,15	3,86
6,55	13,09	19,95	0,81	10,85	12,35
198,38	204,14	263,67	267,90	113,92	0,57
331,77	375,56	433,22	648,40	122,44	0,81
59,37	106,59	168,92	151,63	73,32	0,52
0,47	0,94	1,87	2,03	2,43	0,19
390,21	780,41	1217,99	22,77	150,38	5,61
0,94	1,87	3,74	1,63	3,33	1,05
10,44	9,97	6,86	21,55	5,71	0,73
249,65	297,33	282,37	350,83	129,15	0,63
52,67	105,34	194,48	102,04	80,81	0,21
15,43	21,19	39,27	19,51	18,53	0,05
249,49	285,18	307,93	292,70	135,57	0,54
115,63	225,34	375,25	482,13	147,80	0,69
493,06	386,47	162,07	13,01	93,09	6,16
5,77	11,53	16,21	14,23	9,34	0,34
46,59	46,13	66,70	39,84	31,43	0,21
31,32	32,41	31,17	27,64	17,36	0,37
4,05	8,10	16,21	2,44	8,74	2,58
145,24	177,34	21,82	2,03	19,91	8,80
484,02	651,70	443,81	505,31	201,93	0,60
14,49	28,99	45,50	63,82	21,91	0,66
143,83	287,67	292,34	709,38	131,93	0,81
88,05	129,97	203,21	9,35	86,90	8,29
337,38	521,11	642,66	678,89	241,84	0,64
453,16	603,70	387,71	545,15	182,96	0,66
78,23	156,46	191,36	54,88	87,01	0,59
102,69	205,39	355,92	205,70	140,17	0,32
32,73	52,98	71,06	32,93	34,01	0,03
80,10	89,76	102,23	32,93	49,24	0,50

ANEXO 2 – EXEMPLO DE ENTRADAS E SAÍDAS DO EXPERIMENTO 2

Na tabela abaixo são apresentados exemplos das entradas utilizadas no treinamento do Experimento 2, do valor real vendido, das previsões realizadas pela RNA e do percentual de erro MAPE para cada previsão. Os dados foram alterados por um fator de multiplicação com o objetivo de resguardar a segurança das informações da empresa onde o trabalho foi realizado.

Tabela A: Entradas e saídas do Experimento 2.

90 a 61 dias	60 a 31 dias	30 a 1 dias	Qtd. Vendida	Previsão Exp. 2	MAPE Exp. 2
9,35	9,35	7,48	5,28	7,33	0,39
740,52	2993,87	1413,72	358,96	565,15	0,57
1,87	5,61	1,87	0,41	2,89	6,11
2410,43	908,82	740,52	1336,24	347,71	0,74
39,27	7,48	3,74	1,22	6,86	4,63
286,11	420,75	31,79	2,85	34,81	11,23
104,72	93,50	28,05	24,39	35,69	0,46
1,87	37,40	99,11	15,04	27,87	0,85
37,40	29,92	54,23	53,66	31,20	0,42
52,36	1,87	22,44	47,16	11,36	0,76
16,83	16,83	24,31	20,73	16,04	0,23
149,60	72,93	112,20	16,26	68,95	3,24
16,83	65,45	76,67	14,23	40,01	1,81
20,57	13,09	11,22	11,38	11,02	0,03
14,96	357,17	18,70	163,02	32,57	0,80
99,11	132,77	336,60	133,34	128,75	0,03
35,53	56,10	22,44	15,04	24,53	0,63
84,15	7,48	100,98	27,24	33,20	0,22
9,35	5,61	5,61	39,43	5,64	0,86
9,35	11,22	14,96	12,60	10,40	0,17
336,60	5,61	26,18	15,04	21,02	0,40
9,35	35,53	5,61	0,81	7,89	8,71
63,58	510,51	192,61	267,90	129,00	0,52
422,62	383,35	441,32	648,40	219,20	0,66
31,79	115,94	359,04	151,63	106,80	0,30
1,87	1,87	1,87	2,03	2,10	0,03
2047,65	91,63	112,20	22,77	91,09	3,00
1,87	5,61	3,74	1,63	3,72	1,29
7,48	5,61	7,48	21,55	6,11	0,72
93,50	228,14	508,64	350,83	178,58	0,49
84,15	390,83	46,75	102,04	61,54	0,40
41,14	28,05	48,62	19,51	29,62	0,52
170,17	441,32	310,42	292,70	175,11	0,40
203,83	579,70	342,21	482,13	197,52	0,59
37,40	41,14	407,66	13,01	86,21	5,63
20,57	7,48	7,48	14,23	7,95	0,44
100,98	84,15	14,96	39,84	25,79	0,35
11,22	46,75	35,53	27,64	23,92	0,13
5,61	5,61	37,40	2,44	12,08	3,95
22,44	11,22	31,79	2,03	17,01	7,37
499,29	289,85	493,68	505,31	222,66	0,56
29,92	54,23	41,14	63,82	31,26	0,51
72,93	9,35	787,27	709,38	86,90	0,88
7,48	387,09	213,18	9,35	76,02	7,13
592,79	244,97	1019,15	678,89	314,03	0,54
325,38	499,29	289,85	545,15	182,32	0,67
293,59	220,66	52,36	54,88	64,02	0,17
415,14	183,26	469,37	205,70	194,56	0,05
76,67	29,92	106,59	32,93	48,24	0,46
276,76	18,70	11,22	32,93	18,01	0,45

ANEXO 3 – EXEMPLO DE ENTRADAS E SAÍDAS DO EXPERIMENTO 3

Na tabela abaixo são apresentados exemplos dos dados utilizados no treinamento do Experimento 3, que são os mesmos do Experimento 2, porém, há a adição do mês em que a previsão foi realizada. Os dados foram alterados por um fator de multiplicação com o objetivo de resguardar a segurança das informações da empresa onde o trabalho foi realizado.

Tabela A: Entradas e saídas do Experimento 3.

Mês	90 a 61 dias	60 a 31 dias	30 a 1 dias	Qtd. Vendida	Previsão Exp. 3	MAPE Exp. 3
Agosto	9,35	9,35	7,48	5,28	6,00	0,14
Setembro	740,52	2993,87	1413,72	358,96	542,63	0,51
Julho	1,87	5,61	1,87	0,41	3,47	7,53
Julho	2410,43	908,82	740,52	1336,24	322,06	0,76
Junho	39,27	7,48	3,74	1,22	7,94	5,51
Agosto	286,11	420,75	31,79	2,85	42,73	14,02
Outubro	104,72	93,50	28,05	24,39	42,79	0,75
Novembro	1,87	37,40	99,11	15,04	26,33	0,75
Agosto	37,40	29,92	54,23	53,66	29,85	0,44
Julho	52,36	1,87	22,44	47,16	12,10	0,74
Outubro	16,83	16,83	24,31	20,73	15,64	0,25
Novembro	149,60	72,93	112,20	16,26	55,69	2,42
Janeiro	16,83	65,45	76,67	14,23	36,00	1,53
Junho	20,57	13,09	11,22	11,38	12,46	0,09
Janeiro	14,96	357,17	18,70	163,02	43,23	0,73
Fevereiro	99,11	132,77	336,60	133,34	125,44	0,06
Dezembro	35,53	56,10	22,44	15,04	17,97	0,20
Novembro	84,15	7,48	100,98	27,24	29,32	0,08
Janeiro	9,35	5,61	5,61	39,43	5,86	0,85
Mai	9,35	11,22	14,96	12,60	13,18	0,05
Fevereiro	336,60	5,61	26,18	15,04	19,68	0,31
Novembro	9,35	35,53	5,61	0,81	7,56	8,30
Outubro	63,58	510,51	192,61	267,90	142,04	0,47
Agosto	422,62	383,35	441,32	648,40	221,47	0,66
Dezembro	31,79	115,94	359,04	151,63	85,88	0,43
Setembro	1,87	1,87	1,87	2,03	2,32	0,14
Março	2047,65	91,63	112,20	22,77	142,16	5,24
Junho	1,87	5,61	3,74	1,63	4,40	1,71
Setembro	7,48	5,61	7,48	21,55	6,46	0,70
Abril	93,50	228,14	508,64	350,83	246,12	0,30
Setembro	84,15	390,83	46,75	102,04	46,01	0,55
Outubro	41,14	28,05	48,62	19,51	28,51	0,46
Janeiro	170,17	441,32	310,42	292,70	191,88	0,34
Julho	203,83	579,70	342,21	482,13	169,75	0,65
Outubro	37,40	41,14	407,66	13,01	62,10	3,77
Abril	20,57	7,48	7,48	14,23	11,15	0,22
Setembro	100,98	84,15	14,96	39,84	20,65	0,48
Janeiro	11,22	46,75	35,53	27,64	22,99	0,17
Agosto	5,61	5,61	37,40	2,44	10,63	3,36
Julho	22,44	11,22	31,79	2,03	18,12	7,92
Março	499,29	289,85	493,68	505,31	303,01	0,40
Setembro	29,92	54,23	41,14	63,82	31,67	0,50
Fevereiro	72,93	9,35	787,27	709,38	55,67	0,92
Janeiro	7,48	387,09	213,18	9,35	94,59	9,12
Junho	592,79	244,97	1019,15	678,89	361,08	0,47
Fevereiro	325,38	499,29	289,85	545,15	205,79	0,62
Setembro	293,59	220,66	52,36	54,88	53,59	0,02
Outubro	415,14	183,26	469,37	205,70	176,84	0,14
Agosto	76,67	29,92	106,59	32,93	47,44	0,44
Outubro	276,76	18,70	11,22	32,93	21,45	0,35

ANEXO 4 – ERRO DA REDE NOS EXPERIMENTOS

Na tabela abaixo são apresentados os erros obtidos no treinamento para cada um dos três experimentos, em cada rodada do método *10-fold cross validation*.

Tabela A: Erros obtidos nos experimentos realizados

Rodada	Erro Experimento 1	Erro Experimento 2	Erro Experimento 3
1	0,0204206816976926	0,0187547163461583	0,0177463132525417
2	0,0207733252014266	0,0188286744414124	0,0179778165422184
3	0,0205473775649199	0,0186991639149524	0,0177133831968239
4	0,0208348677290424	0,0191400441108826	0,0180662594665706
5	0,0203005557259510	0,0182899429585431	0,0175168248833322
6	0,0206238489354487	0,0187034133543839	0,0175598949645721
7	0,0207099856662761	0,0187680731759050	0,0177641768600725
8	0,0234544245818700	0,0213541544561052	0,0203279403437441
9	0,0206097901914535	0,0185763708180468	0,0174196762694052
10	0,0208201925441742	0,0184400745660130	0,0171287617445300