

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO
Especialização em Engenharia de Software

Monografia

**Análise Preditiva de Vendas em E-commerce para Tomada de Decisão Estratégica
utilizando Redes Neurais Artificiais**

Luciana Shikano Sato

Projeto desenvolvido durante a especialização em
Engenharia de Software apresentado à Pontifícia
Católica de São Paulo.

Orientador: Prof. Dr. Fernando Antônio de Castro
Giorno

Dezembro de 2015

Sumário

1 Introdução	1
1.1 Objetivos	2
1.2 Contribuição	2
1.3 Método de Trabalho	2
1.4 Organização do Trabalho	3
2 Revisão Bibliográfica	5
2.1 Fundamentos: Redes Neurais Artificiais	5
2.2 Trabalhos Correlatos	17
2.3 Justificativa de utilização das redes neurais artificiais	31
2.4 Limitações	32
3 Formulação de Modelos para Análise Preditiva em E-commerce	33
3.1 Etapas para Formulação de Modelos	33
3.2 Implementação	40
3.3 Análise de Desempenho	47
3.4 Outros experimentos	49
4 Conclusões	50
4.1 Resultados	50
4.2 Sugestões para Trabalhos Futuros	51
Referências Bibliográficas	52
Webgrafia	53

Lista de Figuras

Figura 1: Função de ativação Degrau (limiar)	7
Figura 2: Função de ativação Linear	7
Figura 3: Função de ativação Sigmoidal	8
Figura 4: Função de ativação Tangente Hiperbólica.....	8
Figura 5: Modelo Perceptron	9
Figura 6: Solução de problema linearmente separável.....	9
Figura 7: Funções lógica AND, OR e XOR.....	10
Figura 8: Rede MultiLayer FeedForward.....	11
Figura 9: Rede Feedback.....	12
Figura 10: Rede de Hopfield	13
Figura 11: Funcionamento do algoritmo Back-Propagation	16
Figura 12: Etapas para a exploração dos Modelos de Previsão	20
Figura 13: Arquitetura da pesquisa	27
Figura 14: Obtenção da base de dados	34
Figura 15: Exemplo de RNA para previsão de vendas	39
Figura 16: Representação da arquitetura no MatLab	41
Figura 17: Definição dos exemplos de treinamento, validação e teste no MatLab.....	41
Figura 18: Gráfico de análise de desempenho	48

Lista de Tabelas

Tabela 1: Informações da base de dados	35
Tabela 2: Intervalo de vendas da amostra	35
Tabela 3: Representação da variável Semana	36
Tabela 4: Representação da variável Feriados comerciais / Período promocional	37
Tabela 5: Representação da variável Ação Promocional	37
Tabela 6: Representação da variável Ações de Marketing	37
Tabela 7: Valores de entrada de produto 25182.....	38
Tabela 8: Resultado do treinamento na arquitetura 43-10-4	41
Tabela 9: Resultado do treinamento na arquitetura 43-20-4	42
Tabela 10: Resultado do treinamento na arquitetura 43-30-4	42
Tabela 11: Resultado do treinamento na arquitetura 43-40-4	43
Tabela 12: Resultado do treinamento na arquitetura 43-50-4	43
Tabela 13: Resultado do treinamento na arquitetura 43-60-4	44
Tabela 14: Resultado do treinamento na arquitetura 43-70-4	44
Tabela 15: Resultado do treinamento na arquitetura 43-80-4	45
Tabela 16: Resultado do treinamento na arquitetura 43-90-4	45
Tabela 17: Resultado do treinamento na arquitetura 43-100-4	46
Tabela 18: Resultado do treinamento na arquitetura 43-110-4	46
Tabela 19: Resultado do treinamento na arquitetura 43-120-4	46
Tabela 20: Resultado do treinamento na arquitetura 43-130-4	47
Tabela 21: Resultado do treinamento na arquitetura 43-140-4	47
Tabela 22: Resultado comparativo entre as arquiteturas.....	48
Tabela 23: Comparação entre quantidade de variáveis entre os estudos	49

Lista de Abreviaturas e Siglas

<i>B2B</i>	<i>Business-to-consumer</i>
<i>BPN</i>	<i>Back Propagation Network</i>
<i>GUI</i>	<i>Graphical User Interface</i>
<i>IA</i>	<i>Inteligência Artificial</i>
<i>IDS</i>	<i>Interior Desire System</i>
<i>MCP</i>	<i>Modelo de McCulloch e Pitts</i>
<i>MSE</i>	<i>Mean Squared Error</i>
<i>RNAs</i>	<i>Redes Neurais Artificiais</i>

1 Introdução

O comércio eletrônico (e-commerce) é um setor que vem apresentando uma evolução constante. A praticidade, a comodidade e a economia são quesitos que atraem o consumidor. Eventos promocionais como o “Black Friday”, Dia das Crianças e o Natal movimentam o setor e conseqüentemente o faturamento também aumenta. A crescente popularização da banda larga residencial e móvel e a facilidade na compra também justifica o aumento nas vendas em e-commerce.

Devido a esse crescimento das vendas em e-commerce, as empresas necessitam de um maior planejamento e organização para atender os seus clientes de forma cada vez mais rápida e eficiente. Neste cenário, as empresas tomam decisões estratégicas seguindo tendências baseadas na realidade em que estão inseridas. Algumas empresas apóiam-se em estatísticas de vendas dos produtos para elaborar seu estoque, outras, de forma passiva, esperam a confirmação de um pedido do cliente para buscar o produto em questão, ou trabalham com estoque reduzido. Estes dois cenários geram inconvenientes tanto para a empresa quanto para o cliente.

No primeiro cenário, a empresa infla seu estoque de produtos que podem ficar parados. Isso é muito negativo para a empresa pois com o estoque parado não gera lucro para a empresa, além dos gastos com o armazenamento.

O segundo cenário interfere na decisão de compra do cliente. Se o produto em questão não está disponível em estoque para venda, a empresa necessita contactar o fornecedor para realizar a compra e entregar o produto para o cliente. Esse processo de buscar o produto com o fornecedor resulta em um tempo de entrega maior para o cliente.

Diversas variáveis podem interferir na decisão de compra do cliente como valor do frete, promoção, forma de pagamento, forma de devolução do produto, troca do produto e tempo de entrega entre outros.

Com o avanço da capacidade de armazenamento dos dados e com a utilização de técnicas de data mining é possível extrair padrões de comportamento referentes as

compras dos consumidores por meio do histórico das transações armazenadas. Essas informações são relevantes para a empresa que consegue analisar o perfil de seus clientes. Para a área de marketing, essa técnica possibilita a elaboração de campanhas promocionais direcionadas para perfis diferentes de seus clientes.

Além disso, a elaboração de modelos preditivos podem ajudar eficientemente para a identificação do cenário de vendas futuro, possibilitando o planejamento controlado de ações das diversas áreas da empresa. Para a construção de um sistema de previsão de vendas, as Redes Neurais Artificiais são frequentemente utilizadas por apresentarem um resultado representativo.

1.1 Objetivos

Este trabalho tem como objetivo analisar como as Redes Neurais Artificiais (RNAs) podem auxiliar a área de compras de um e-commerce na tomada de decisão estratégica, permitindo prever a venda de um determinado produto em um e-commerce.

Este trabalho visa também abordar a topologia de multicamadas nas RNAs, os algoritmos de aprendizagem das RNAs com foco na tomada de decisão estratégica e apresentar a elaboração de uma RNA configurada com parâmetros relativos a um e-commerce .

1.2 Contribuição

Como contribuição é apresentado um estudo para a concepção inicial de um sistema baseado em redes neurais artificiais capaz de analisar a previsão de vendas. A proposta da pesquisa apresenta-se com maior ênfase em resultados comparativos de desempenho entre as redes neurais apresentadas.

1.3 Método de Trabalho

A pesquisa sobre análise preditiva de vendas em um e-commerce foi iniciada por meio de uma levantamento bibliográfico que inclui análise de trabalhos

correlatos e o estudo de fundamentos e paradigmas relacionados à Inteligência Artificial. Nesta atividade foi feita uma pesquisa sobre os modelos neurais existentes, as topologias e os algoritmos de treinamento.

O trabalho seguiu com um estudo mais aprofundado sobre os algoritmos de treinamentos. Nesta fase, foi possível definir algumas arquiteturas de RNA utilizadas na pesquisa.

O próximo passo da pesquisa foi a criação e elaboração de uma base de teste que simule os dados de um e-commerce.

Com as definições das RNAs e com uma base de teste preparada, foi possível a elaboração de algumas redes neurais com diferentes algoritmos de treinamento. Nesta fase, a criação e o treinamento das RNAs foram realizados utilizando o software MatLab.

Finalizado os testes com as redes neurais e os diversos algoritmos foi efetuada uma análise de desempenho das redes criadas com o cenário de um e-commerce.

Por fim, foram descritos os resultados obtidos e a rede que apresentou o melhor desempenho.

1.4 Organização do Trabalho

No capítulo 2, “Revisão Bibliográfica”, são apresentados os fundamentos das Redes Neurais Artificiais como modelos neurais, topologias e aprendizado. É apresentado também os trabalhos correlatos, a justificativa de utilização das redes neurais artificiais com multicamadas e as limitações encontradas.

No capítulo 3, “Formulação de Modelos para Análise Preditiva em E-commerce” são descritas as etapas para a formulação dos modelos, a implementação, a análise do desempenho das arquiteturas e outros experimentos realizados.

No capítulo 4, “Conclusão”, são descritos os resultados obtidos, as restrições para o projeto de previsão de vendas, bem como sugestões para a continuidade desse trabalho.

2 Revisão Bibliográfica

2.1 Fundamentos: Redes Neurais Artificiais

A Inteligência Artificial (IA) é uma área da Ciência da Computação que visa atribuir comportamentos humanos inteligentes em sistemas. Segundo Russel & Norvig (2009), as definições da IA podem ser divididas em quatro categorias:

1. Sistemas que pensam como humanos
2. Sistemas que pensam racionalmente
3. Sistemas que agem como humanos
4. Sistemas que agem racionalmente

A IA é uma ciência fundamentada em diversas áreas de conhecimentos, como Linguística, Psicologia, Ciência Cognitiva, Biologia, Matemática, Engenharia, Planejamento e Computação. Por ser multidisciplinar é capaz de solucionar problemas em vários campos. Pode ser aplicada em sistemas autônomos, jogos, diagnósticos, predição, classificação e robótica, entre outras áreas.

Existem diversas técnicas estudadas na IA, cada uma apresenta modelos característicos e peculiares, como sistemas de agentes, algoritmos genéticos, árvores de decisão, sistemas especialistas, computação evolutiva, lógica fuzzy e redes neurais artificiais.

As RNAs estão inseridas no modelo conexionista dentro da Inteligência Artificial.

Braga et al. (2000) definem RNAs como:

“RNAs são sistemas paralelos distribuídos compostos por unidades de processamento simples (nodos) que calculam determinadas funções matemáticas (normalmente não-lineares). Tais unidades são dispostas em uma ou mais camadas e interligadas por um grande número de conexões, geralmente unidirecionais. Na maioria dos modelos estas conexões estão associadas a pesos,

os quais armazenam o conhecimento representado no modelo e servem para ponderar a entrada recebida por cada neurônio da rede. O funcionamento destas redes é inspirado em uma estrutura física concebida pela natureza: o cérebro humano.”

Modelo de McCulloch-Pitts (MCP)

O primeiro modelo de neurônio foi proposto por Warren McCulloch e Walter Pitts em 1943. McCulloch era um psiquiatra e neuroanatomista e Pitts era um matemático. O modelo apresentado por McCulloch e Pitts era uma simplificação do que se sabia a respeito do neurônio biológico.

“Um neurônio biológico dispara quando a soma dos impulsos que ele recebe ultrapassa o seu limiar de excitação (threshold). O corpo do neurônio, por sua vez, é emulado por um mecanismo simples que faz a soma dos valores $x_i w_i$ recebidos pelo neurônio (soma ponderada) e decide se o neurônio deve ou não disparar (saída igual a 1 ou a 0) comparando a soma obtida ao limiar ou threshold do neurônio.” (Braga et al., 2000, p. 8).

No modelo de McCulloch e Pitts a ativação do neurônio é obtida utilizando a função de ativação que ativa (1) ou não (0) a saída, ou seja, um neurônio binário. Neste caso as respostas do neurônio é dada de forma binária. No modelo proposto originalmente a função de ativação é dada pela função de limiar.

Segundo Braga et al. (2000), as redes neurais artificiais se diferem das redes neurais biológicas no aspecto físico e as redes neurais artificiais buscam reproduzir as funções das redes biológicas com o intuito de implementar o comportamento básico.

Função de Ativação

A função de ativação é responsável por ativar o neurônio ou não. Se os valores de entrada estiverem corretos, esta função ativará o neurônio, caso contrário o neurônio ficará desativado.

No modelo McCulloch e Pitts a função de ativação do modelo proposto é a função degrau (0 ou 1). Este modelo deu origem a utilização de outras funções de ativação, como: função linear, função sigmoideal, função tangente hiperbólica.

As funções abaixo são alguns exemplos de função de ativação:

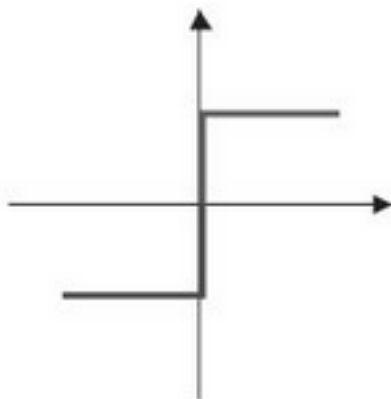


Figura 1: Função de ativação Degrau (limiar) (Simas, 2013)

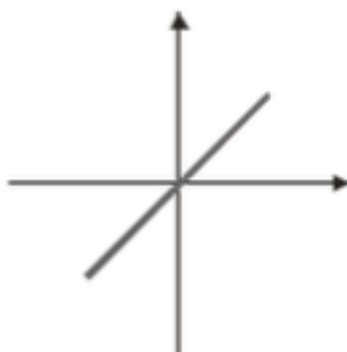


Figura 2: Função de ativação Linear (Simas, 2013)

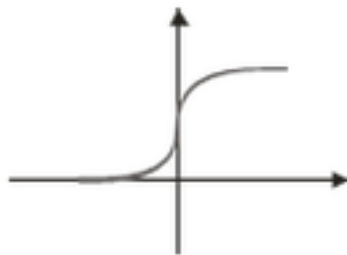


Figura 3: Função de ativação Sigmoidal (Simas, 2013)

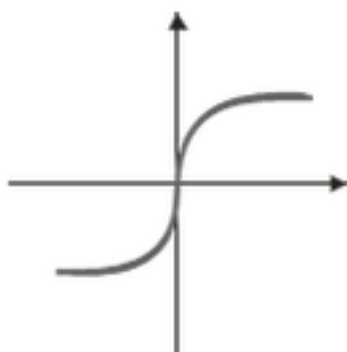


Figura 4: Função de ativação Tangente Hiperbólica (Simas, 2013)

Perceptron

O perceptron foi proposto por Rosenblatt em 1958. Este modelo é composto por uma estrutura de rede apresentando neurônios e uma regra de aprendizado. A topologia possui neurônios na camada de entrada, um nível intermediário e um nível de saída, ou seja, uma estrutura de três níveis embora conhecida como perceptron de uma única camada.

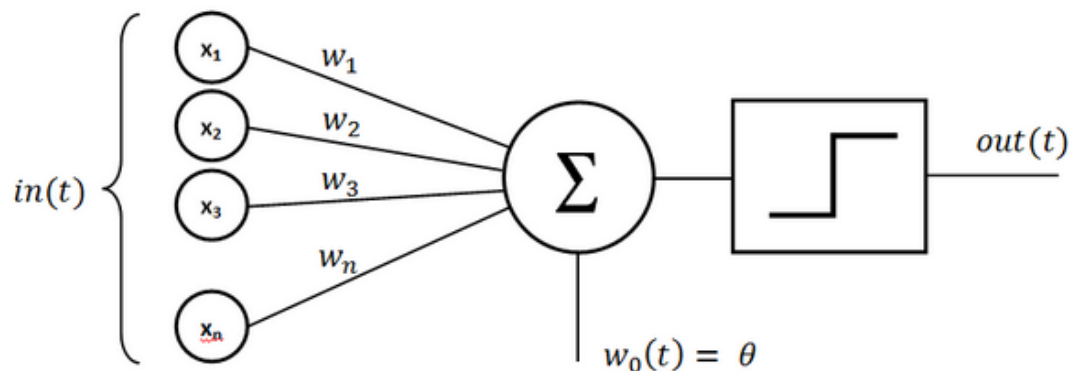


Figura 5: Modelo Perceptron (Wang, 2014)

O modelo proposto por Rosenblatt é uma solução de classificação e reconhecimento de padrões. Neste modelo, os padrões são linearmente separáveis ou seja, o perceptron tem a capacidade de separar padrões apenas de maneira linear, como visto no figura 6.

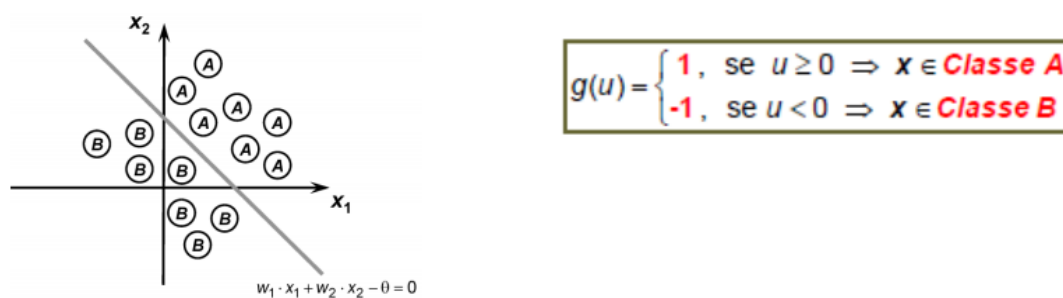


Figura 6: Solução de problema linearmente separável (Watanabe)

Uma limitação do modelo perceptron é sua incapacidade de resolver problemas não lineares. O perceptron não é capaz de solucionar, por exemplo, a função lógica XOR (Ou exclusivo). Após alguns anos da apresentação do modelo perceptron, o modelo de múltiplas camadas (Multilayer Perceptron) surgiu e foi capaz de solucionar problemas não lineares como a função XOR.

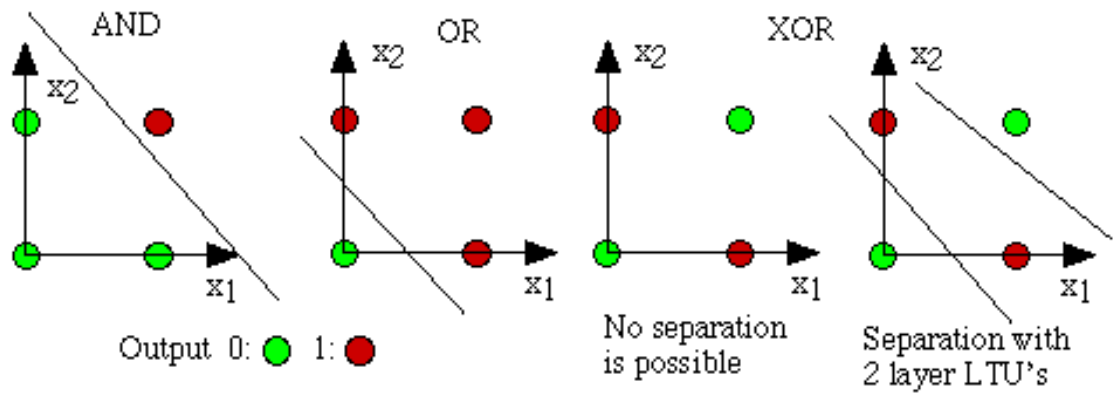


Figura 7: Funções lógica AND, OR e XOR

Topologia das RNAs

A topologia das RNAs podem ser categorizadas pela forma como propagam a informação recebida, ou seja, em Redes Feedforward (propagação para frente) ou Redes Feedback (realimentação).

As redes feedforward apresentam unidades de processamentos (neurônios) que propagam a informação da camada de entrada, passando pela camada intermediária até a camada de saída.

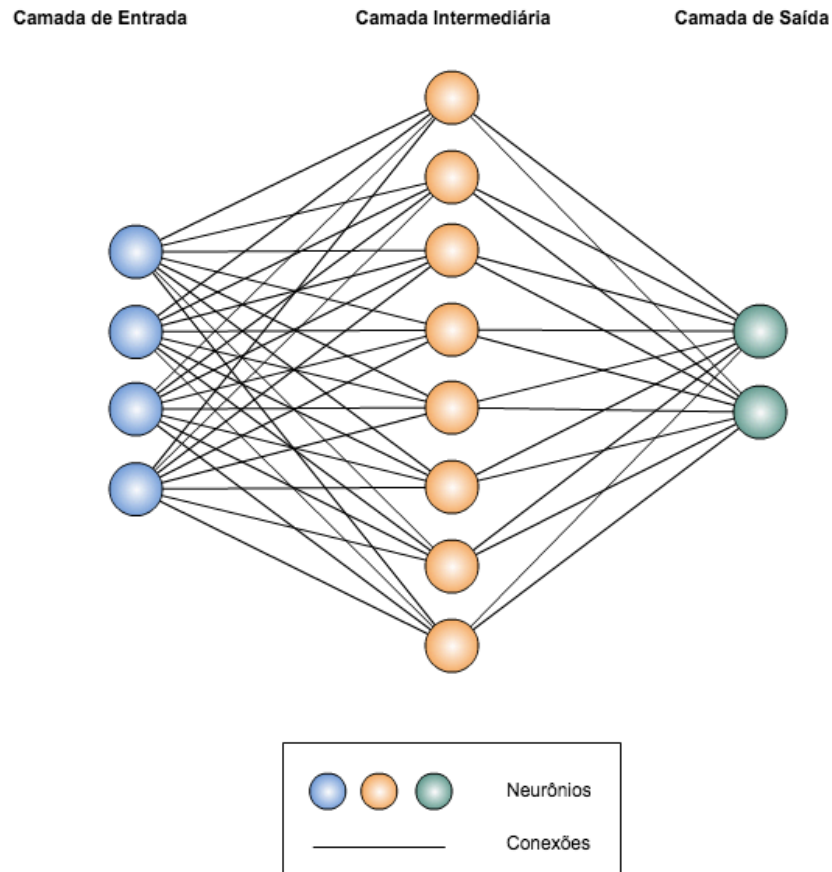


Figura 8: Rede MultiLayer FeedForward

Haykin (2001) define *realimentação* (feedback) como sendo:

“Dizemos que existe em um sistema dinâmico sempre que a saída de um elemento do sistema influencia em parte a entrada aplicada àquele elemento particular, originando assim um ou mais de uma caminho fechado para transmissão de sinais em torno do sistema.” (Haykin, 2001, p.44)

As redes feedback diferenciam das redes feedforward pelo fato de apresentarem ao menos uma recorrência, ou seja, o sinal de saída de um neurônio é a entrada de outro neurônio anterior.

A figura 9 é um exemplo de redes neurais feedback na qual a saída do último neurônio é uma das entradas do neurônio da camada anterior (intermediária).

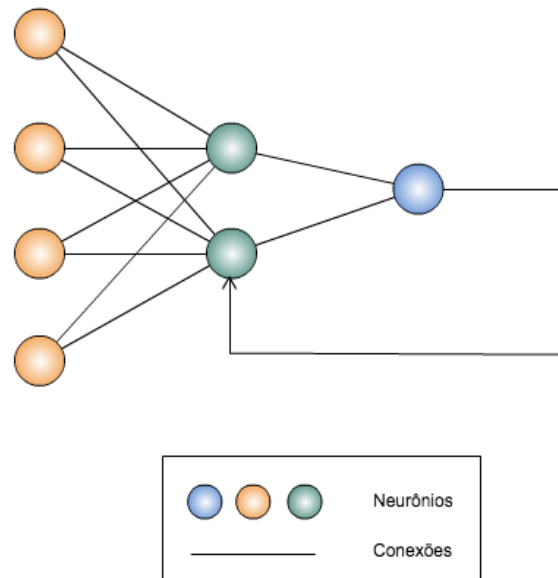


Figura 9: Rede Feedback

A figura 10 é um exemplo de uma rede feedback conhecida como modelo recorrente de Hopfield. Essa rede não apresenta linearidade e as saídas dos neurônios estão associadas às entradas.

Segundo Braga et al. (2000), a recorrência dá ao modelo características temporais que implicam que a resposta da rede dependa sempre do seu estado no intervalo de tempo anterior.

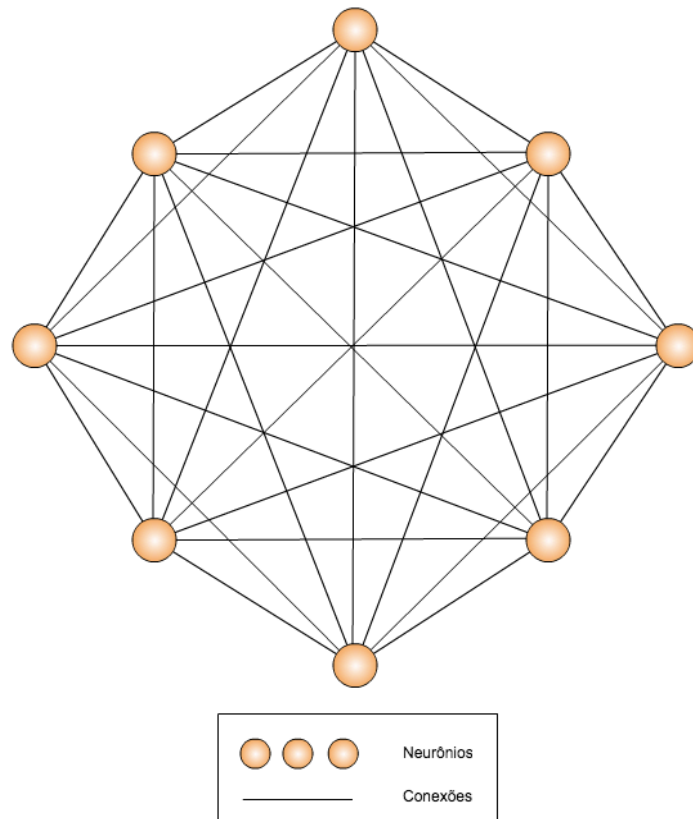


Figura 10: Rede de Hopfield

Aprendizado das RNAs

Redes Neurais Artificiais possuem a capacidade de aprender por exemplos e fazer interpolações e extrapolações do que aprenderam. No aprendizado conexionista, não se procura obter regras como na abordagem simbólica da Inteligência Artificial (IA), mas sim determinar a intensidade de conexões entre neurônios (Braga et al., 2000).

O aprendizado nas RNAs pode ser classificado de duas maneiras: aprendizado supervisionado e aprendizado não-supervisionado.

No *aprendizado supervisionado* as entradas e saídas desejadas para a rede são fornecidas por um supervisor (professor) externo. O objetivo é ajustar os parâmetros da rede, de forma a encontrar uma ligação entre os pares de entrada e saída fornecidos (Braga et al., 2000).

No *aprendizado não-supervisionado* o sistema deve ser auto-organizado podendo adquirir diversas formas estruturais. Uma outra característica é que o aprendizado não apresenta pares de entrada e saída, como no supervisionado. Segundo a definição de Ferneda (2006), a rede processa as entradas e, detectando suas regularidades, tenta progressivamente estabelecer representações internas para codificar características e classificá-las automaticamente (Ferneda, 2006).

Algoritmos de Treinamento

Os algoritmos de treinamento de RNAs multicamadas podem ser classificados em estáticos e dinâmicos.

Os algoritmos estáticos não alteram a estrutura da rede, variando apenas os valores de seus pesos. Os algoritmos dinâmicos podem tanto reduzir quanto aumentar o tamanho da rede (número de camadas, número de nodos nas camadas intermediárias e número de conexões) (Braga et al., 2000).

O algoritmo de treinamento conhecido como Back-Propagation é utilizado em RNAs multicamadas com aprendizado supervisionado. Este algoritmo necessita de pares de entrada e saída desejada para que haja correções de erros e ajustes de pesos quando ocorre o treinamento.

O treinamento deste algoritmo ocorre em duas fases: a forward e a backward. A fase forward inicia com a apresentação dos valores à camada de entrada da RNA. Os sinais de cada neurônio é calculado e são produzidos os valores de saída da RNA. Esses valores são comparados com os valores de saída desejados que resulta em um valor de erro. Se esse valor estiver acima do permitido, ou seja, a rede ainda não aprendeu da forma como se esperava, o erro é retropropagado para camadas anteriores. Essa retropropagação ocorre na fase backward. A partir da camada de saída, os neurônios sofrem ajustes em seus

pesos com a finalidade de reduzir os erros. Esses ajustes são realizados nos neurônios da camada intermediária e nos neurônios da camada de entrada.

Esse processo ocorre até que o erro seja mínimo (dentro do limite estabelecido) ou até atingir a quantidade de épocas (iterações) definidas na configuração inicial RNA.

Braga et al. (2000) apresenta alguns pontos negativos na utilização deste algoritmo. Segundo ele, o principal problema diz respeito à lentidão do algoritmo em superfícies mais complexas. Apesar desse problema, algumas técnicas podem ser usadas para contorná-la. São elas:

- Utilizar taxa de aprendizado decrescente;
- Adicionar nós intermediários;
- Utilizar um termo *momentum*;
- Adicionar ruído aos dados.

Um outro ponto negativo é quando em um determinado ciclo de treinamento, a rede ao invés de melhorar com os ajustes dos pesos, começa a piorar a sua taxa de acertos. Esse problema é conhecido como ocorrência de *overfitting*. Isso demonstra que a RNA memorizou os padrões de treinamento. Para este tipo de problema, Braga et al. (2000) também sugere algumas alternativas, como:

- Encerrar o treinamento mais cedo (quando o erro de validação começa a subir);
- Podar (*pruning*) os pesos da rede.

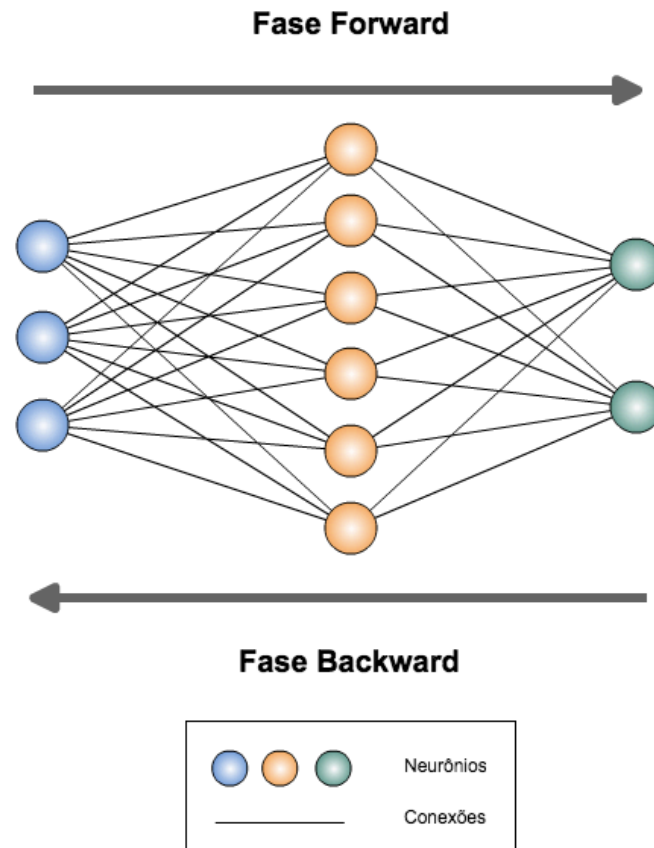


Figura 11: Funcionamento do algoritmo Back-Propagation

Mean Squared Error (MSE)

O erro médio quadrático ou MSE é o erro médio entre a saída da rede neural e os valores desejados. É normalmente usado como indicador de desempenho da rede neural para avaliar seus resultados. Durante o treinamento da rede, o MSE começa com valores altos e decresce conforme a quantidade de épocas aumenta. Ao final do treinamento, o MSE se mantém constante e a rede neural converge.

2.2 Trabalhos Correlatos

Exploração de Dados Atomizados para Previsão de Vendas no Varejo Utilizando Redes Neurais

O trabalho de pesquisa apresentado por Passari (2003) apresenta um intenso estudo no uso de redes neurais na previsão de vendas no varejo. O objetivo central foi a geração de previsões de demanda para a tomada de decisão no varejo a partir de dados históricos de vendas de produtos de uma dada empresa. Há uma preocupação em explorar uma metodologia baseada em técnicas de mineração de dados para a identificação de relacionamentos entre demanda de produtos e suas influências. Este trabalho visa ajudar os gestores de vendas na tomada de decisão de forma mais precisa. “A metodologia explorada busca ser útil, portanto, para tomada de decisões operacionais no varejo – e não estratégicas, de longo prazo” (Passari, 2003).

Para a elaboração dos testes dos modelos de previsão de vendas foram abordados três aspectos: nível de agregação dos dados, seleção de variáveis preditivas e o composto de marketing e a seleção de produtos afins.

Na classificação das técnicas de previsão segundo o nível de agregação dos dados, a previsão de vendas podem variar conforme o nível de agregação dos dados. O nível pode ser individual ou agregado e deve-se considerar o tempo de existência do produto no mercado – novo ou existente – pois esse fator é influente na previsão de vendas. No trabalho de Passari (2003) foi utilizada a classificação de nível individual com os produtos já existentes no mercado, pois segundo o autor, considera a sensibilidade das vendas de um produto às variáveis do composto de marketing.

O segundo aspecto considerado pela pesquisa de Passari (2003) é a seleção de variáveis preditivas e o composto de marketing. Essa escolha de variáveis são fundamentais para a previsão de vendas futuras de um determinado produto. A previsão de vendas foi baseada em compostos de marketing do varejo. O autor

apresenta algumas definições de outros autores sobre a definição de composto de marketing.

Definição de Parente (2000, p.61):

“o conjunto de variáveis que compõem o esforço de marketing do varejista e engloba todos os fatores controláveis que o varejista articula [...] para conquistar as preferências dos consumidores”

Definição de Mason, Mayer e Wilkinson (1993):

“O composto de marketing do varejo pode ser classificado a partir de seis macrofatores chamados seis Ps: mix de produtos, preço, promoção, apresentação, pessoal, ponto-de-venda”.

Baseado nessas definições, Passari (2003) conclui que a fatia de mercado de uma loja depende principalmente das características de seu composto de marketing. Acredita também que o comportamento do consumidor em uma loja é influenciado por três características: o composto de marketing da loja, o composto de marketing dos concorrentes da loja e das variáveis ambientais, ou seja, dada a economia, o clima, a demografia, a época do ano e as tendências do mercado.

A partir da pesquisa desse segundo aspecto, o autor apresenta seu trabalho baseado nos componentes do composto de marketing: produto, preço e promoção e justifica que a escolha desses componentes ocorreu devido a dificuldade de obtenção de dados representativos dos outros componentes, como apresentação, pessoal e pontos-de-venda.

O terceiro aspecto considerado na pesquisa é a seleção de produtos afins – identificação de centro de interesse. Segundo o autor não é viável tratar em um só modelo todos os produtos de uma loja (48.000 itens no caso da pesquisa em questão). Foi utilizada a técnica de análise dos cestos de compras que permite selecionar um conjunto de produtos que possuem correlação em suas vendas.

Neste caso, buscou-se a previsão de demanda dentro desse centro de interesse em particular.

Definido o critério da técnica de análise de cestos de compras, são geradas regras como “se o consumidor compra o produto A, há uma probabilidade P de que ele também compre o produto B”.

Finalizada a apresentação dos três aspectos citados acima, o autor inicia a apresentação de redes neurais aplicadas à previsão, explica o método de aprendizado por retropropagação e também o funcionamento das redes de multicamadas.

O método de pesquisa de Passari (2003) é do tipo experimental ou seja, busca a capacidade preditiva de técnicas empregadas. Essa etapa teve o objetivo em elaborar redes neurais artificiais para a previsão de vendas a fim de prever a demanda individualizada dos produtos. A exploração dos dados, a construção e a avaliação das redes neurais seguiram seis passos.

1. Obtenção de dados
2. Seleção e limpeza
3. Análise de cesto de compras: centros de interesse
4. Tratamento de dados
5. Geração de modelos baseados em redes neurais
6. Comparação com outras modelagens

Passari (2003) demonstra essas etapas com uma figura que define a sequência de etapas apresentada em seu trabalho.

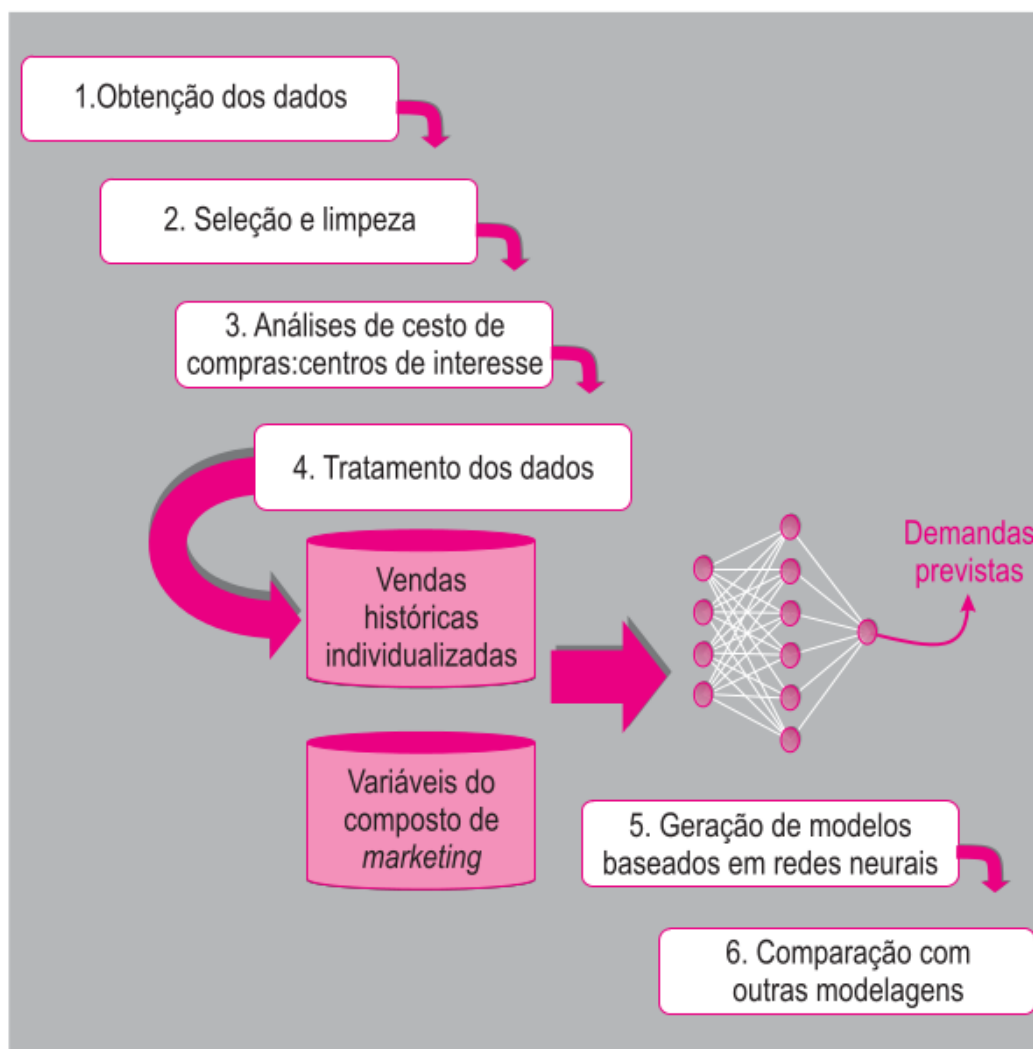


Figura 12: Etapas para a exploração dos Modelos de Previsão (Passari, 2003)

Para as etapas de tratamento dos dados, identificação dos centros de interesse e criação dos modelos de previsão, o autor utilizou algumas ferramentas de software de data mining. Para o tratamento inicial dos dados foi utilizado o SPSS e para a criação das redes neurais foram utilizados o pacote Intelligent Miner da IBM, o Mega-puter Poly Analyst 4.5 e o Statistica Neural Networks.

Na etapa de desenvolvimento e resultados da pesquisa, o autor apresenta algumas características da base de dados que foi explorada. Essa base de dados pertence a uma rede varejista brasileira que comercializa produtos bastantes variados como eletroeletrônicos, utilidades domésticas, móveis, vestuários, brinquedos entre outros. A rede de pontos-de-venda é distribuída em todo território nacional com mais de 100 pontos.

As redes neurais foram criadas e testadas com diversas configurações para analisar o melhor desempenho. O autor elaborou 144 redes neurais diferentes e conclui que as redes que apresentaram o melhor desempenho tinham sofrido normalizações e padrões de dados e também supressão de variáveis correlacionadas. Outra observação feita pelo autor foi que o melhor desempenho foi observado quando utilizou-se vendas passadas de produtos inter-relacionados em um mesmo centro de interesse do consumidor.

Os dados da loja utilizados na pesquisa foram coletados em um período de vendas da empresa de novembro de 1998 a junho de 2002. Foram utilizados dados produtos vendidos em apenas um pontos-de-venda devido ao volume de dados e limitações de tratamento desses dados. Esse pontos-de-venda foi selecionado por representantes da própria empresa segundo seu interesse, ou seja, tamanho da loja, localização, mix de produtos e principalmente a diversidade de itens comercializados neste local.

A partir da escolha da loja foram definidas algumas questões referentes a utilização da base de dados da loja selecionada. Apenas vendas realizadas a prazo, ou seja, 60% das vendas da loja foram utilizadas na base. Nessas vendas haviam informações mais precisa sobre o cliente permitindo análises do comportamento do consumo do cliente. Um outro filtro aplicado à base de dados dessa loja foi a inutilização de vendas no qual haviam inconsistências nos registros como datas de vendas inconsistentes, juros negativos, estoques sem registros, ausência de classificação de produto, entre outras. Esses registros que apresentavam alguma inconsistência foram removidas da base da pesquisa.

Ao final da aplicação de filtros e remoção de dados inconsistentes da base de dados, a base de dados final da pesquisa apresentou 12.560 registros sendo 881 produtos, 558 marcas distintas, periodicidade semanal e vendas somente a prazo.

Na análise dos cestos de compras foi utilizada, segundo o autor, com o propósito de separar um grupo coeso de produtos para minimizar o total de produtos a serem previstos pelo modelo. Segundo Passari (2003) o objetivo era elaborar um

modelo de previsão de vendas de um centro de interesse e não de todos os produtos da empresa em questão. Além disso, essa técnica foi importante para investigar se produtos que fazem parte de um mesmo centro de interesse podem trazer alguma melhora na capacidade preditiva de modelos de previsão baseados em redes neurais.

O autor apresentou duas estratégias na abordagem dos dados: análise por meio de transações, ou seja, buscar quais itens são vendidos em conjunto numa mesma venda e a análise horizontal por cliente na qual exibe quais produtos um mesmo cliente compra ao longo do tempo. Nessa etapa de desenvolvimento foi utilizado os softwares Intelligent Miner 6.1 (IBM, 1999) para a identificação do centro de interesse. Essa identificação foi feita visualmente a partir de gráficos gerados pelo software.

A seleção das variáveis foi baseada na disponibilidade na base de dados sendo que não houve critério conceitual para a sua escolha. As variáveis utilizadas na pesquisa foram:

1. Semana: Variável de tempo, contabilizando as semanas sequencialmente. A contagem inicia-se em 01 de novembro de 1998, considerada como semana 1. A última semana, que se inicia no dia 26 de maio de 2002, é a semana 187.
2. Preço médio semanal: Média de preço praticado por produto, durante a semana correspondente.
3. Juros médios semanais: Taxa de juros média praticada por produto, durante a semana correspondente.
4. Prazo: Porcentagem de vendas realizadas a prazo na semana anterior.
5. id_produto, produto, linha, setor, id_marca: Variáveis de identificação do produto.

6. Classe, classe_distribuição: A empresa adota duas classificações distintas para um mesmo produto: uma baseada no giro do produto, e outra baseada no tipo de compra/distribuição a que ele está submetido. Ambas, entretanto, identificam parâmetros de alta, média ou baixa venda, venda sob encomenda, venda sem mostruário, venda direta do fornecedor e produto fora de linha.

7. Comissão: Indica se o produto entra na comissão dos vendedores na semana correspondente.

8. Estoque, Estoque_dias, Reposição_dias, Mostruário, Grade: Representam a disponibilidade do produto, seja em estoque, seja em mostruário, seja em prazo de entrega/reposição.

9. Campanha, Campanha_agressiva, Veiculações: Variáveis que medem o esforço promocional efetuado na semana. Para esforços corporativos, todos os produtos receberam o mesmo valor para a variável Campanha no período correspondente. As campanhas agressivas referem-se a apenas alguns grupos de produtos, enquanto Veiculações mede a exposição das campanhas na mídia.

10. Variáveis com lag de tempo: Foram testadas variáveis com lag de tempo para as condições em que se detectaram correlações horizontais.

A princípio, Passari (2003) utilizou os dados em suas escalas originais para detectar os produtos inter-relacionados e encontrar os centros de interesse do consumidor. Em seguida, os dados originais foram transformados para aproximar sua distribuição da normal e melhorar a assimilação dos dados por modelos quantitativos.

Com relação as variáveis de entradas das redes neurais artificiais, o autor concluiu que de todas as variáveis incorporadas ao modelo apenas a informação sobre o total do produto vendido na semana anterior interferiu significativamente na capacidade preditiva da rede.

Uma outra etapa do desenvolvimento da pesquisa foi avaliar a capacidade preditiva do modelo baseado em redes neurais. A base de dados foi dividida em dois grupos:

1. Grupo de criação do modelo, constituído de registros de treinamento e validação. Nesse grupo foram utilizados registros de novembro de 1998 a maio de 2001, totalizando 135 semanas e sempre prevendo uma semana à frente.
2. Grupo de avaliação externo que simula o modelo de previsão proposto com dados desconhecidos. Os registros utilizados foram subsequente a semana 135 ou seja, de maio de 2001 a junho de 2002 totalizando 3140 casos.

Nesta etapa o autor define, basicamente, os dados que foram utilizados para treinar as redes neurais propostas e a base utilizada a validar e testar os modelos de redes na neurais elaborados na pesquisa.

Passari (2003), em sua pesquisa, elaborou 141 redes neurais e fez um trabalho detalhado de analisar o comportamento e melhor resultado entre elas. Em seu estudo verificou que a rede neural com melhor desempenho apresentado foi a rede 108 com MSE de 1,7457. A taxa média de acertos para previsão de uma semana à frente é de 41,21%.

As principais características da rede 108 que obteve o melhor desempenho são:

1. Variáveis de entrada da rede neural: Id_produto, Comissão, Setor, Linha, Id_marca, Classe, Classe_distribuição, Estoque_logP, Estoque_dias-2P, Reposição_diasP, Semana, Preço_potP, JurosMédiosSemanaisP, PprazoP, Mostuário_logP, Lag1, Vendas_linhaP
2. Variável prevista: vendas_log
3. Quantidade de nós: 7
4. Número de camadas: 3

5. Erro Quadrado: 0,7164

6. MSE: 1,7457

Para finalizar o estudo e exploração das redes neurais, Passari (2003) comparou o resultado obtido das redes neurais com outros dois modelos de previsão: análise de regressão e Método naïve.

Para o modelo de regressão foram utilizadas as mesmas variáveis do modelo de rede 108, a rede que apresentou o melhor resultado. O resultado obtido com a técnica de regressão linear utilizando o grupo de validação foi de um MSE de 49,1%, superior ao resultado obtido utilizando as redes neurais que foi de 41,21%.

O autor verificou que o desempenho das redes neurais foi superior a 54% em relação ao Método naïve considerando o erro quadrado médio (RMSE) e superior a 64% considerando o erro médio percentual (MAPE).

Em suas conclusões, Passari (2003) afirma que o desempenho das redes foi favorável na previsão de vendas de produtos no nível individual, superando outros métodos como regressão linear e Método naïve. Conclui que as redes neurais são capazes de tratar os dados de diferentes produtos e que os dados agregados apresentam relação entre os produtos. O histórico de vendas foi de fundamental importância para a previsão das vendas. Afirma também que a taxa de previsão foi baixa considerando os 41% de acerto para previsão uma semana à frente, ou seja, alguns desvios influenciaram no aprendizado das redes neurais como qualidade dos dados disponíveis, baixo número de registros por produto, poucas informações sobre as promoções, movimentações macroeconômicas, falta de informações sobre as concorrentes e inconsistências de dados.

Integrating Web Mining and Neural Network for Personalized e-Commerce Automatic Service

O trabalho de Chou et al. (2009) ilustra a mineração de dados na web baseada no comportamento de navegação do usuário de forma a descobrir padrões de navegação nos sites. Tem o objetivo de prever os futuros clientes em potencial. Utiliza o modelo de Redes Neurais Artificiais com multicamadas para buscar tais resultados na pesquisa. *“Our method is based on the customer’s on-line navigation behaviors by analyzing their navigation patterns through web mining and constructing artificial neural networks to predict potential customers’ need in the future”* (Chou et al., 2009).

Chou et al. (2009) iniciam o trabalho apresentando uma introdução sobre e-commerces e vendas no modelo B2B (business-to-consumer). Na sequência, citam a importância de estratégias de marketing para as empresas e descrevem a valor sobre o conhecimento do produto (*product knowledge*). Afirmam que o *product knowledge* é importante tanto para o comprador quanto para o vendedor e também o fabricante.

Os autores apresentam definições tanto para data mining (mineração de dados) quanto para web mining. O data mining refere-se a extração de conhecimento de uma base de dados. Já o web mining também refere-se a extração de conhecimento porém baseado em dados da web. Como exemplo, os autores citam o comportamento de navegação de estudantes que utilizam o sistema de e-learning. Essa análise de comportamento pode ajudar na melhoria de interface da plataforma para melhorar a qualidade do serviço prestado.

Na sequência é apresentada uma introdução sobre as redes neurais e a teoria sobre o funcionamento das redes neurais de multicamadas.

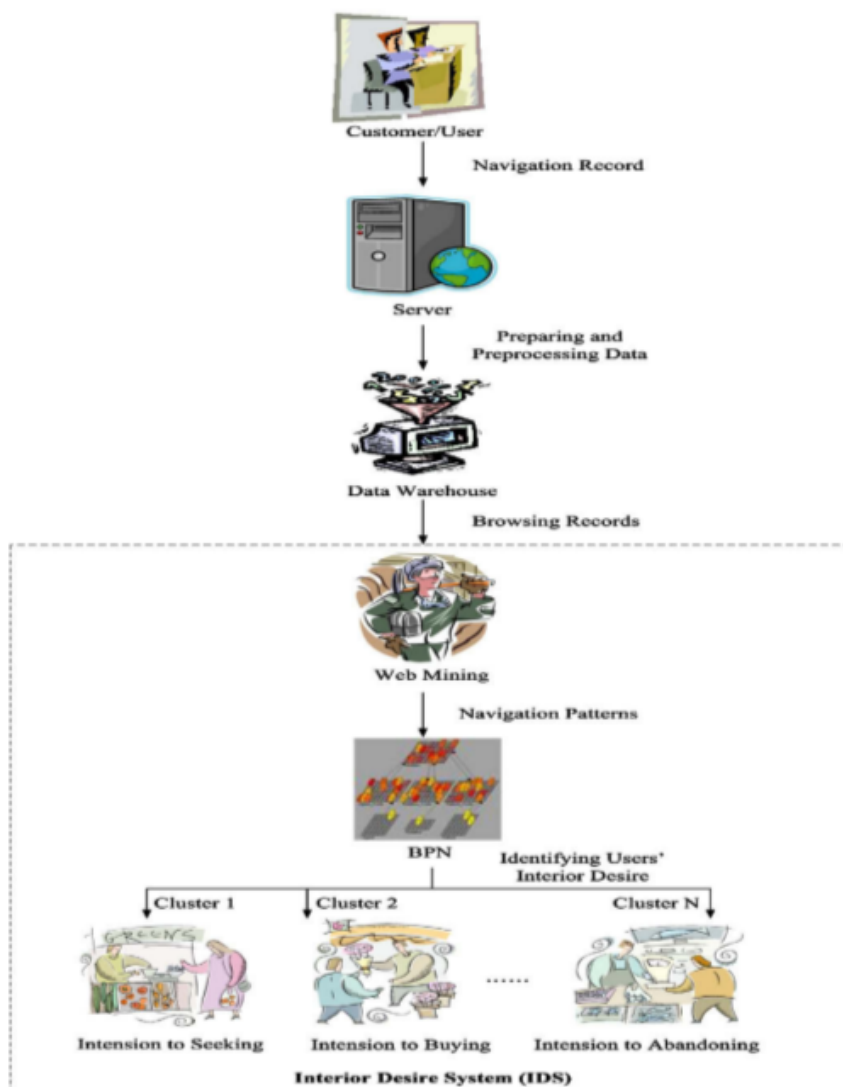


Figura 13: Arquitetura da pesquisa (Chou et al., 2009)

A metodologia de pesquisa deste trabalho apresenta três fases incluindo seis passos. Os autores acreditam que se os usuários possuem padrões de navegação parecidos, logo eles podem ter interesses semelhantes para determinados produtos. Eles propõem um sistema denominado *interior desire system (IDS)* para avaliar o conhecimento do produto de potencial interesse pelo cliente e propor os produtos apropriados para os mesmos.

A primeira fase denominada Modelagem de Dados é composta por dois passos:

Passo 1 - Gravação da navegação

Essa etapa consiste em gravar toda a navegação feita pelo usuário, ou seja, quando um usuário inicia uma sessão em um website o IDS invocaria o gravador de navegação para iniciar a gravação do comportamento e toda a navegação feito pelo mesmo. Dessa forma, seria possível iniciar a elaboração da base de dados para a realização do data mining.

Passo 2 – Preparação dos dados e pré processamento

Essa etapa incluem a remoção de dados que apresentam ruídos deixando apenas na base de dados informações que serão relevantes para a modelagem do sistema. Por exemplo, arquivos que apresentam a extensão gif ou jpeg podem ser removidas pois não serão úteis.

A fase dois é denominada Modelagem baseada em conteúdo e apresenta dois passos:

Passo 3 – Armazenamento de dados

Após a fase 1 é possível coletar as informações referentes a navegação do usuário e todo o fluxo realizado em um website gravados em um arquivo log. Nessa etapa é possível armazenar todos os cliques realizados pelo usuário até a finalização de uma compra, por exemplo.

Passo 4 – Web Mining

Nesse etapa é feita a análise do caminho e análise de padrões sequenciais. A descoberta de padrões sequenciais em logs de acesso do servidor web permite que as empresas encontrem padrões de visitas do usuário e ajuda na segmentação de publicidades direcionando a grupos específicos de usuários com base nesses padrões.

A fase três é definida como Modelagem de filtragem colaborativa e apresenta dois passos:

Passo 5 – Implementação do BPN (Back Propagation Network)

Nesse passo o objetivo é classificar os usuários em grupos com padrões de comportamento similares. Esses comportamentos de usuários são usados para realizar a previsão de comportamentos futuros dos usuários e identificar um potencial interesse em produtos ou até a intenção de compras.

Passo 6 – Sistema de recomendação

A partir da elaboração dos cinco passos citados anteriormente, esse passo refere-se a construção de um sistema de recomendação para usuários. Esse sistema identifica os padrões comportamentais dos usuários e recomenda produtos para futuras compras.

Esta pesquisa utilizou a base de dados de um site do Reino Unido que vendia produtos relacionados a cuidados da pele. Os autores utilizaram a base de dados referente ao período de 01 de dezembro de 2004 a 31 de dezembro de 2005.

Para a elaboração das redes neurais foi utilizado 629,821 registros de navegação dos usuários.

Os autores criaram algumas redes neurais com configurações diferentes baseados em quantidade de épocas utilizadas para o treinamento das redes e quantidade de neurônios na camadas intermediárias distintas. Para a quantidade de épocas no treinamento foram utilizadas 50, 100, 200 e 300. Já a quantidade de neurônios na camadas intermediárias foram 8, 12, 16, 24 e 32.

A rede neural que apresentou o melhor resultado foi a rede configurada com 100 épocas e 8 neurônios na camada intermediária.

A pesquisa concluiu que tanto o comportamento quanto a intenção do usuário podem ser identificadas utilizando os métodos dessa pesquisa e é possível classificar padrões e recomendar produtos baseados em comportamentos

anteriores. Afirmam também que é possível construir um sistema de recomendações utilizando técnicas de web mining e técnicas de BPN.

Aplicação de Redes Neurais na Tomada de Decisão no Mercado de Ações

O trabalho de Gambogi & Costa (2013) apresenta um sistema de trading para tomada de decisão de compra e venda do índice Standard & Poor's 500, na modalidade seguidor de tendência utilizando redes neurais artificiais multicamadas com propagação para frente dentro de um período de 5 anos.

Os autores iniciam o trabalho apresentando uma introdução sobre o mercado financeiro e a teoria sobre as redes neurais artificiais. O sistema de trading tem o objetivo de operar uma carteira que simula o índice Standard & Poor's 500 com a finalidade de auferir lucros de forma consistente e contínua. As operações de compra e venda da carteira são realizadas somente nos finais dos últimos pregões de cada semana pelos preços de fechamento do índice.

A metodologia aplicada no trabalho seguiu as seguintes sequencias:

1. Configuração da arquitetura da rede neural

Nesta etapa foram definidas as entradas da rede neural, o algoritmo de treinamento utilizado foi o Bayesian Regularization do MATLAB, uma camada intermediária com funções de ativação tangentes hiperbólicas (tansig) e um neurônio na camada de saída. Esse neurônio representa o preço de fechamento semanal estimado. O período da base de dados utilizada nesse trabalho foi de 06/04/1990 e 29/06/2012 totalizando 1160 semanas. Dessas 1160 semanas, 900 semanas foram utilizadas para treinamento, validação e testes e as outras 260 semanas restantes foram utilizadas para o teste da rede neural proposta.

2. Seleção de atributos

A escolha dos atributos foi inspirada no trabalho de Myungssok (2005) que empregou os fechamentos semanais nas 4 semanas anteriores à semana da estimativa da rede neural e o índice técnico MACD (moving average convergence and divergence) na semana imediatamente anterior.

3. Configurações testadas para integrar o sistema de trading

Os autores deste trabalho elaboraram diversas arquiteturas de redes neurais modificando a quantidade de neurônios na camada intermediária. A partir da elaboração, treinamento e testes das diversas redes neurais, as arquiteturas candidatas a integrar o sistema de trading foram pré-selecionadas pelos autores.

4. Critérios de escolha das configurações pré-selecionadas

O critério de escolha das redes neurais levou em consideração a taxa de acerto nas previsões das oscilações semanais do índice acima de 60% na amostra de teste e que estava geralmente associada a um maior retorno anual nominal do sistema de trading no período de verificação.

Na escolha das configurações capazes de apresentar bons retornos no período de verificação foram geradas 1000 redes neurais.

Utilizando o MATLAB, dessas 1000 redes neurais foram escolhidas 5 redes que apresentaram a menor taxa de erros na amostra de teste.

Os resultados obtidos mostraram que a rede neural que apresentou melhor resultado foi de 75,28% de acertos das operações.

2.3 Justificativa de utilização das redes neurais artificiais

De acordo com Russel & Norvig (2009), as redes neurais são capazes de aprender tarefas complexas. Como exemplo, os autores apresentam um estudo

de caso de reconhecimento de dígitos escritos a mão. Este exemplo é um problema importante e complexo na qual a solução apresentada pelos autores é utilizando redes neurais artificiais de multicamadas.

A pesquisa de Passari (2003) apresenta algumas justificativas para a utilização de redes neurais, como: robustas e tolerantes a falhas, não-lineares, comportam-se bem com problemas onde o conhecimento a respeito é de difícil estruturação, capazes de aproximar qualquer função contínua ou série temporal e possuem boa capacidade de generalização.

2.4 Limitações

Segundo Braga et al (2000) existe uma dificuldade e deficiência no treinamento das redes neurais utilizando o algoritmo de backpropagation. O principal problema é com relação a lentidão do algoritmo para superfícies mais complexas. O algoritmo pode convergir para mínimos locais, ou seja, pontos na superfície de erro que apresenta a solução estável, embora não seja a saída correta.

3 Formulação de Modelos para Análise Preditiva em E-commerce

3.1 Etapas para Formulação de Modelos

A formulação dos modelos foi dividida em cinco etapas desde a geração de uma base de conhecimento para a pesquisa até a análise dos resultados obtidos. Nessas etapas foi utilizado alguns conceitos de Data Mining.

Na primeira etapa foi elaborada uma base de dados para utilização na pesquisa. Essa base de dados foi gerada randomicamente, ou seja, foi criado um script que simulava compras de usuários de um e-commerce. A estrutura da base de dados encontra-se no Anexo I.

Em seguida houve a escolha de produtos para análise e predição das vendas. O critério de escolha dos produtos foi feita a partir de produtos mais vendidos dentro de um período de vendas. Dentre os produtos mais vendidos foi escolhida uma amostra de 198 produtos.

Após a seleção desses produtos foi iniciada a fase de tratamento dos dados para a criação das RNAs. O tratamento da base de dados consistiu em transformar os valores da amostra em valores binários relacionando valores de entradas e valores de saída esperados.

Ao final da fase de tratamento de dados foi iniciada a criação de RNAs para treinamento, testes e análise. Para todas as RNAs criadas foi utilizado o software MatLab.

A última fase de análise e comparação entre as RNAs foi baseada nos resultados obtidos pelo software. Foram criadas 14 redes para o teste e análise. Todas as redes apresentavam a mesma quantidade de neurônios na camada de entrada e saída e o que variava foi a quantidades de neurônios na camada intermediária.

As fases do método de pesquisa contemplaram:

1. Obtenção de dados
2. Seleção dos produtos para o estudo
3. Tratamento dos dados selecionados
4. Criação das RNAs
5. Análise e comparação entre RNAs

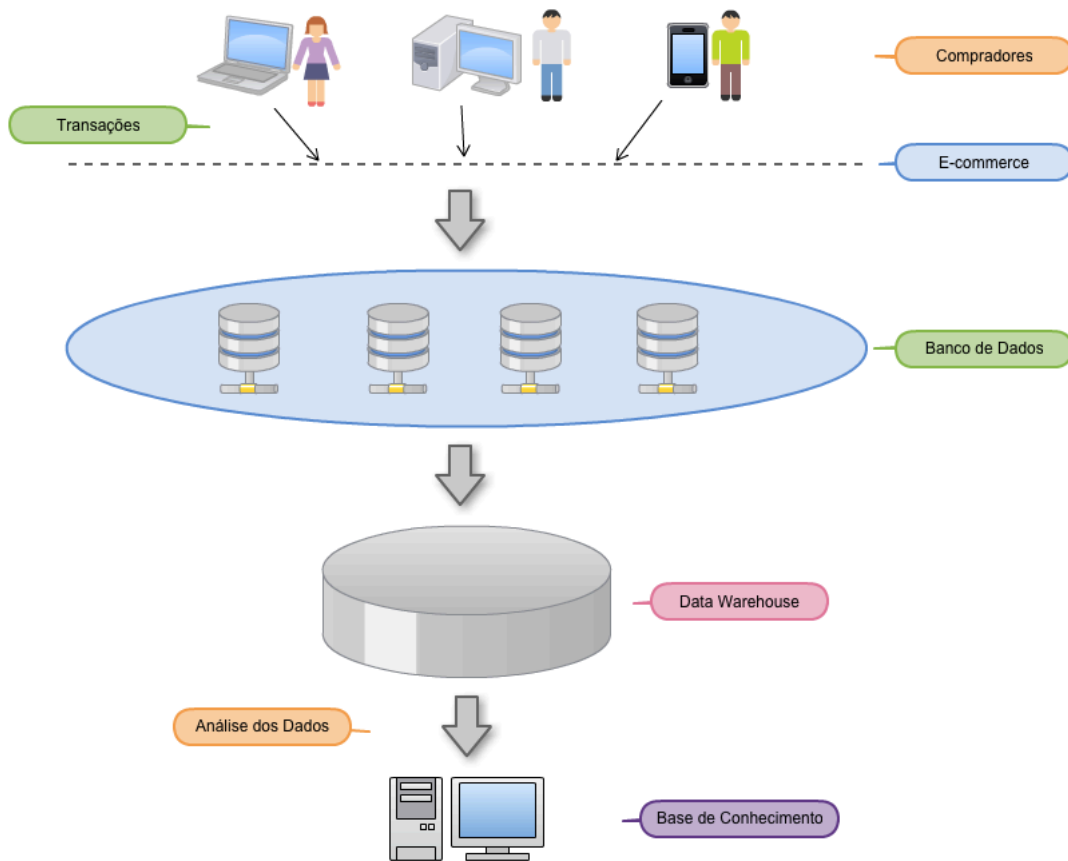


Figura 14: Obtenção da base de dados

Obtenção dos dados

A base de dados utilizada foi criada para fins da pesquisa. Para isso foram criadas as tabelas de categorias, subcategorias, produtos e vendas.

	Quantidade
Produtos Vendidos	591.774
Vendas	143.831
Produtos	36.512
Categorias	6
Subcategorias	11

Tabela 1: Informações da base de dados

Seleção dos itens de estudo

O período de análise deste estudo foi escolhido entre as datas 12/10/2014 à 31/12/2014.

Os produtos mais vendidos neste período foram selecionados para o estudo, portanto, a amostra apresentou 198 produtos. O critério de seleção foi baseado nas vendas consolidadas entre o mais vendido, cujo valor foi de 21 até o limite de 16.

	Amostra (198 produtos)
Mais Vendido da amostra	21
Menos Vendido da amostra	16

Tabela 2: Intervalo de vendas da amostra

Tratamento dos dados selecionados

Após a escolha dos produtos foram definidas as variáveis de entrada utilizadas para a previsão de vendas.

As variáveis de entrada definidas foram:

- SKU (Identificação do produto)
- Semana
- Quantidade de vendas na semana anterior
- Feriados comerciais / Período promocional
- Preço médio semanal
- Tempo médio de entrega do pedido
- Ações promocionais
- Ações de Marketing

A variável semana foi representada da seguinte forma:

Período	Semana
12/10/2014 à 18/01/2014	42
19/10/2014 à 25/10/2014	43
26/10/2014 à 01/11/2014	44
02/11/2014 à 08/11/2014	45
09/11/2014 à 15/11/2014	46
16/11/2014 à 22/11/2014	47
23/11/2014 à 29/11/2014	48
30/11/2014 à 06/12/2014	49
07/12/2014 à 13/12/2014	50
14/12/2014 à 20/12/2014	51
21/12/2014 à 27/12/2014	52

Tabela 3: Representação da variável Semana

A variável feriados comerciais / período promocional é referente a datas na qual o comércio tem um notável aumento nas vendas. Para o período de análise foram consideradas os seguintes feriados e datas promocionais:

Feriado/ Data Promocional	Período	Semana
Black Friday	23/11/2014 a 29/11/2014	40
Período do Natal	14/12/2014 a 20/12/2014	51
Período do Natal	21/12/2014 a 27/12/2014	52

Tabela 4: Representação da variável Feriados comerciais / Período promocional

As variáveis Preço médio semanal e Tempo médio de entrega do pedido foram calculadas semanalmente e foi feito arredondamento dos número decimais.

A variável Ações promocionais indica se na semana houve alguma promoção com o produto, como por exemplo, desconto progressivos, desconto por categoria entre outros. Essa variável foi representada pela seguinte forma:

	Sim	Não
Ações Promocionais	1	0

Tabela 5: Representação da variável Ação Promocional

A variável Ações de Marketing indica se na semana houve alguma propaganda ou divulgação em mídias de um determinado produto. A representação dela é:

	Sim	Não
Ações de Marketing	1	0

Tabela 6: Representação da variável Ações de Marketing

Como o estudo baseia-se em RNAs de multicamadas com aprendizado supervisionado, cada valor de entrada da RNA deve apresentar um valor de saída esperado. Neste caso, o valor de saída foi a quantidade de vendas do produto semanalmente.

Variáveis de Entrada		Valores de Entrada										
SKU		25182	25182	25182	25182	25182	25182	25182	25182	25182	25182	25182
Semana		42	43	44	45	46	47	48	49	50	51	52
Vendas semana anterior		0	3	3	1	1	1	1	2	3	1	3
Feriados / Período Promocional		0	0	0	0	0	0	1	0	0	1	1
Preço médio semanal		33,74	34,32	33,85	32,81	31,76	34,9	34,73	33,62	32,46	34,2	34,9
Tempo médio de entrega		7	8	1	9	3	4	10	4	1	3	4
Ações Promocionais		1	0	0	0	0	1	0	0	0	0	1
Ações de Marketing		0	0	0	0	1	1	0	1	0	1	0
Saída Esperada		Valores de Saída										
Quantidade de vendas		3	3	1	1	1	1	2	3	1	3	2

Tabela 7: Valores de entrada de produto 25182

A figura abaixo representa a relação entre as variáveis de entrada com os neurônios da camada de entrada da RNA e o neurônio de saída representado pela quantidade de vendas. Este neurônio de saída representa o valor de previsão de vendas resultado do treinamento da RNA.

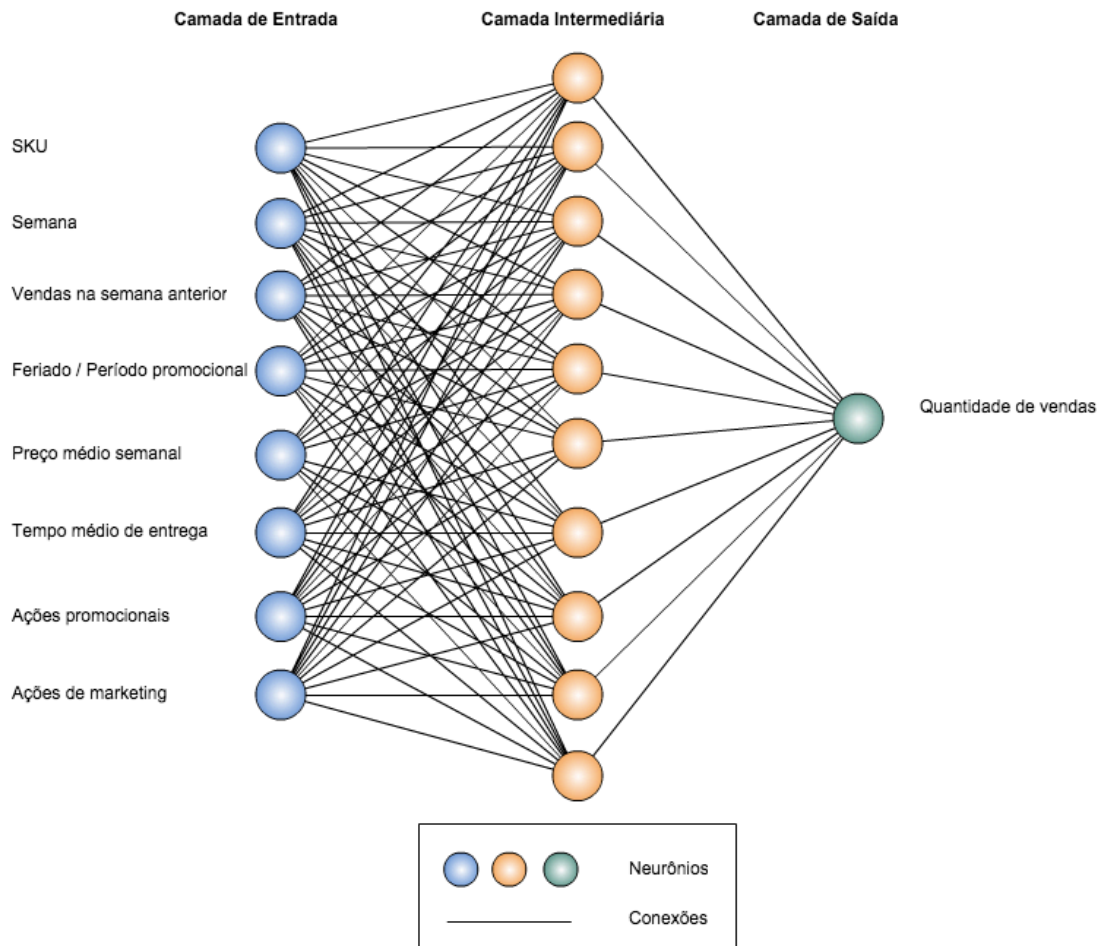


Figura 15: Exemplo de RNA para previsão de vendas

A amostra contém 198 produtos, sendo 11 semanas analisadas. Portanto, a entrada de dados da RNA resultou em uma matriz 8×2178 , sendo:

- Número de variáveis = 8
- $198 \text{ produtos} \times 11 \text{ semanas} = 2178$

Como a RNA da pesquisa apresenta apenas um neurônio de saída, a matriz de resultado esperado foi representada por um vetor 1×2178 .

Tanto os valores de entrada quanto o de saída foram convertidos para binário resultando em um matriz de entrada 43×2178 e a saída em uma matriz 4×2178 .

Dessa forma a arquitetura da RNA do estudo apresenta 43 neurônios de entrada e 4 neurônios de saída.

Essa conversão foi utilizada pois dessa forma as RNAs apresentavam uma taxa de acertos maior.

3.2 Implementação

A fase de implementação iniciou-se com a criação de algumas RNAs para análise de desempenho cujo objetivo foi encontrar a arquitetura da RNA que apresentasse o melhor desempenho.

A criação, treinamento e testes de todas as redes implementadas foram feitas utilizando o *nnstart*, uma ferramenta GUI do MatLab.

A representação das arquiteturas foi descrita no formato Arquitetura x-y-z, sendo:

- x: quantidade de neurônios na camada de entrada
- y: quantidade de neurônios na camada intermediária
- z: quantidade de neurônios na camada de saída

Abaixo seguem as configurações das redes e suas principais características.

Nesse estudo foram criadas redes usando o *Pattern recognition and classification*, do *nnstart*, com as seguintes características:

Arquitetura 43-10-4

A figura abaixo mostra a forma como o MatLab representa as arquiteturas criadas pela ferramenta.

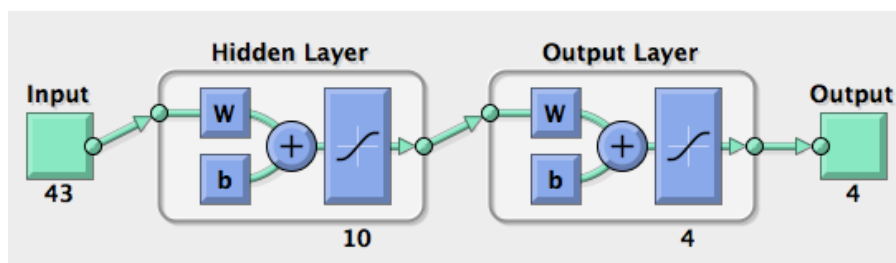


Figura 16: Representação da arquitetura no MatLab

O treinamento da rede com esta arquitetura resultou:

	Exemplos	MSE	%E	% Previsão
Treinamento	1524	9.66288e-2	38,39764	61,60236
Validação	327	1.08021e-1	42,50764	57,49236
Teste	327	1.06808e-1	42,81345	57,18655

Tabela 8: Resultado do treinamento na arquitetura 43-10-4

A definição da quantidade de exemplos a serem treinados, a quantidade de exemplos para a validação e para testes pode ser definido no MatLab. Por padrão é definido 70% das entradas para treinamento, 15% para validação e teste.

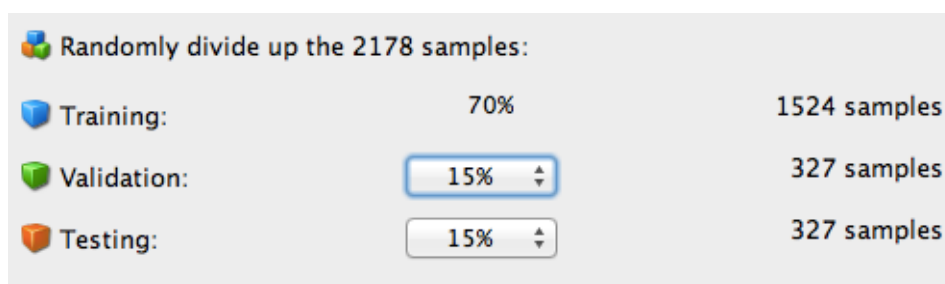


Figura 17: Definição dos exemplos de treinamento, validação e teste no MatLab

Portanto, na arquitetura utilizando 10 neurônios na camada intermediária resultou em uma taxa de 57,19% de acerto para a previsão de vendas.

Arquitetura 43-20-4

Para este estudo foi criada uma rede com 20 camadas intermediárias. O resultado apresentado foi:

	Exemplos	MSE	%E	% Previsão
Treinamento	1524	9.44674e-2	37,53280	62,46720
Validação	327	1.08925e-1	41,28440	58,71560
Teste	327	1.04602e-1	40,51987	59,48013

Tabela 9: Resultado do treinamento na arquitetura 43-20-4

Para esta arquitetura, o resultado foi melhor que o anterior, com 10 neurônios na camada intermediária. O resultado apresentado foi de 59,48% de acertos na previsão de vendas.

Arquitetura 43-30-4

Para este estudo foi criada uma rede com 30 camadas intermediárias. O resultado apresentado foi:

	Exemplos	MSE	%E	% Previsão
Treinamento	1524	9.04280e-2	35,53149	64,46851
Validação	327	1.19995e-1	44,49541	55,50459
Teste	327	1.18642e-1	43,88379	56,11621

Tabela 10: Resultado do treinamento na arquitetura 43-30-4

Para a arquitetura com 30 neurônios na camada intermediária, a taxa de foi abaixo das demais anteriores, porém a taxa de acertos no treinamento foi o mais alto até então, 64,47% de acertos.

Arquitetura 43-40-4

Para este estudo foi criada uma rede com 30 camadas intermediárias. O resultado apresentado foi:

	Exemplos	MSE	%E	% Previsão
Treinamento	1524	9,67475e-2	38,28740	61,7126
Validação	327	1,15221e-1	43,88379	56,11621
Teste	327	1,06570e-1	40,21406	59,78594

Tabela 11: Resultado do treinamento na arquitetura 43-40-4

Com uma arquitetura de 40 camadas intermediárias, o resultado para a previsão de vendas foi de 59,78% na análise dos testes.

Arquitetura 43-50-4

Para este estudo foi criada uma rede com 50 camadas intermediárias. O resultado apresentado foi:

	Exemplos	MSE	%E	% Previsão
Treinamento	1524	1,04180e-2	40,15748	59,84252
Validação	327	1,19996e-1	43,73088	56,26912
Teste	327	1,18105e-1	39,90825	60,09175

Tabela 12: Resultado do treinamento na arquitetura 43-50-4

Com 50 neurônios na camada intermediária, o resultado apresentado foi de 60,10% de acertos na previsão de vendas.

Arquitetura 43-60-4

Para este estudo foi criada uma rede com 60 camadas intermediárias. O resultado apresentado foi:

	Exemplos	MSE	%E	% Previsão
Treinamento	1524	1,00612e-2	39,17322	60,82678
Validação	327	1,11272e-1	43,73088	56,26912
Teste	327	1,23480e-1	45,41284	54,58716

Tabela 13: Resultado do treinamento na arquitetura 43-60-4

O resultado para esta arquitetura foi de 60,83% de previsão de acertos, porém a taxa de acertos no teste foi de apenas 54,59%.

Arquitetura 43-70-4

Para este estudo foi criada uma rede com 70 camadas intermediárias. O resultado apresentado foi:

	Exemplos	MSE	%E	% Previsão
Treinamento	1524	1,00534e-1	39,33727	60,66273
Validação	327	1,11405e-1	42,81345	57,18655
Teste	327	1,34162e-1	47,40061	52,59939

Tabela 14: Resultado do treinamento na arquitetura 43-70-4

Para a arquitetura de 70 camadas intermediárias, a taxa de acertos no treinamento foi de 60,67% e no teste foi de 52,60%.

Arquitetura 43-80-4

Para este estudo foi criada uma rede com 80 camadas intermediárias. O resultado apresentado foi:

	Exemplos	MSE	%E	% Previsão
Treinamento	1524	1,00955e-2	38,48425	61,51575
Validação	327	1,19689e-1	45,87155	54,12845
Teste	327	1,11889e-1	43,11926	56,88074

Tabela 15: Resultado do treinamento na arquitetura 43-30-4

Para uma arquitetura com 80 neurônios na camada intermediária, o resultado apresentado foi de 56,88% de acertos na previsão de vendas.

Arquitetura 43-90-4

Para este estudo foi criada uma rede com 90 camadas intermediárias. O resultado apresentado foi:

	Exemplos	MSE	%E	% Previsão
Treinamento	1524	1,07741e-2	42,02755	57,97245
Validação	327	1,15745e-1	40,21406	59,78595
Teste	327	1,13669e-1	43,88379	56,11621

Tabela 16: Resultado do treinamento na arquitetura 43-90-4

Com a arquitetura de 90 camadas intermediárias, a taxa de acertos de treinamento e testes foi de, respectivamente, 57,97% e 56,12%. A variação da taxa de acertos entre eles foi pequena.

Arquitetura 43-100-4

Para este estudo foi criada uma rede com 100 camadas intermediárias. O resultado apresentado foi:

	Exemplos	MSE	%E	% Previsão
Treinamento	1524	1,05847e-2	40,38713	59,61287
Validação	327	1,09063e-1	42,20183	57,79817
Teste	327	1,11660e-1	44,64831	55,35169

Tabela 17: Resultado do treinamento na arquitetura 43-100-4

Para uma arquitetura com 100 neurônios na camada intermediária, o resultado foi abaixo dos demais, apenas 55,35% de acertos.

Arquitetura 43-110-4

Para este estudo foi criada uma rede com 110 camadas intermediárias. O resultado apresentado foi:

	Exemplos	MSE	%E	% Previsão
Treinamento	1524	1,03545e-1	38,94356	61,05644
Validação	327	1,17905e-1	44,64831	55,35169
Teste	327	1,08730e-1	43,11926	56,88074

Tabela 18: Resultado do treinamento na arquitetura 43-110-4

A taxa de acertos para esta arquitetura no treinamento foi de 61,06%, porém a taxa no teste foi de apenas 56,88%.

Arquitetura 43-120-4

Para este estudo foi criada uma rede com 120 camadas intermediárias. O resultado apresentado foi:

	Exemplos	MSE	%E	% Previsão
Treinamento	1524	1,06967e-1	39,66535	60,33465
Validação	327	1,13656e-1	42,35474	57,64526
Teste	327	1,14696e-1	45,71865	54,28135

Tabela 19: Resultado do treinamento na arquitetura 43-120-4

Para esta arquitetura a taxa de acertos foi de 60,33% e no teste foi de apenas 54,28%.

Arquitetura 43-130-4

Para este estudo foi criada uma rede com 130 camadas intermediárias. O resultado apresentado foi:

	Exemplos	MSE	%E	% Previsão
Treinamento	1524	1,07814e-1	39,69816	60,30184
Validação	327	1,08335e-1	42,20183	57,79817
Teste	327	1,04822e-1	42,96636	57,03364

Tabela 20: Resultado do treinamento na arquitetura 43-130-4

A taxa de acertos nesta arquitetura foi de 60,30% e no treinamento foi de 57,03%.

Arquitetura 43-140-4

Para este estudo foi criada uma rede com 140 camadas intermediárias. O resultado apresentado foi:

	Exemplos	MSE	%E	% Previsão
Treinamento	1524	1,07813e-1	40,25590	59,74410
Validação	327	1,28942e-1	43,88379	56,11621
Teste	327	1,23790e-1	46,94189	53,05811

Tabela 21: Resultado do treinamento na arquitetura 43-140-4

Com esta arquitetura de 140 camadas intermediárias a taxa de acertos foi de 59,74% e apenas 53,06% no processo de teste.

3.3 Análise de Desempenho

Notou-se que a melhor arquitetura encontrada foi a 43-50-4 que apresentou uma taxa de acertos de 60,09%. As outras duas arquiteturas que alcançaram um taxa

de acertos próxima à arquitetura 43-50-4 foram as arquiteturas com 40 e 20 neurônios na camada intermediária.

Percebeu-se também que a partir de 50 neurônios na camada intermediária, a taxa de acertos variou entre 52 e 57% e o aumento dos neurônios na camada intermediária não significou aumento na taxa de acertos.

Neurônios	10	20	30	40	50	60	70	80	90	100	110	120	130	140
Treinem.	61,6	62,47	64,47	61,71	59,84	60,83	60,66	61,52	57,97	59,61	61,06	60,33	60,3	59,74
Validação	57,49	58,72	55,5	56,12	56,27	56,27	57,19	54,13	59,79	57,8	55,35	57,65	57,8	56,12
Teste	57,19	59,48	56,12	59,79	60,09	54,59	52,6	56,88	56,12	55,35	56,88	54,28	57,03	53,06

Tabela 22: Resultado comparativo entre as arquiteturas

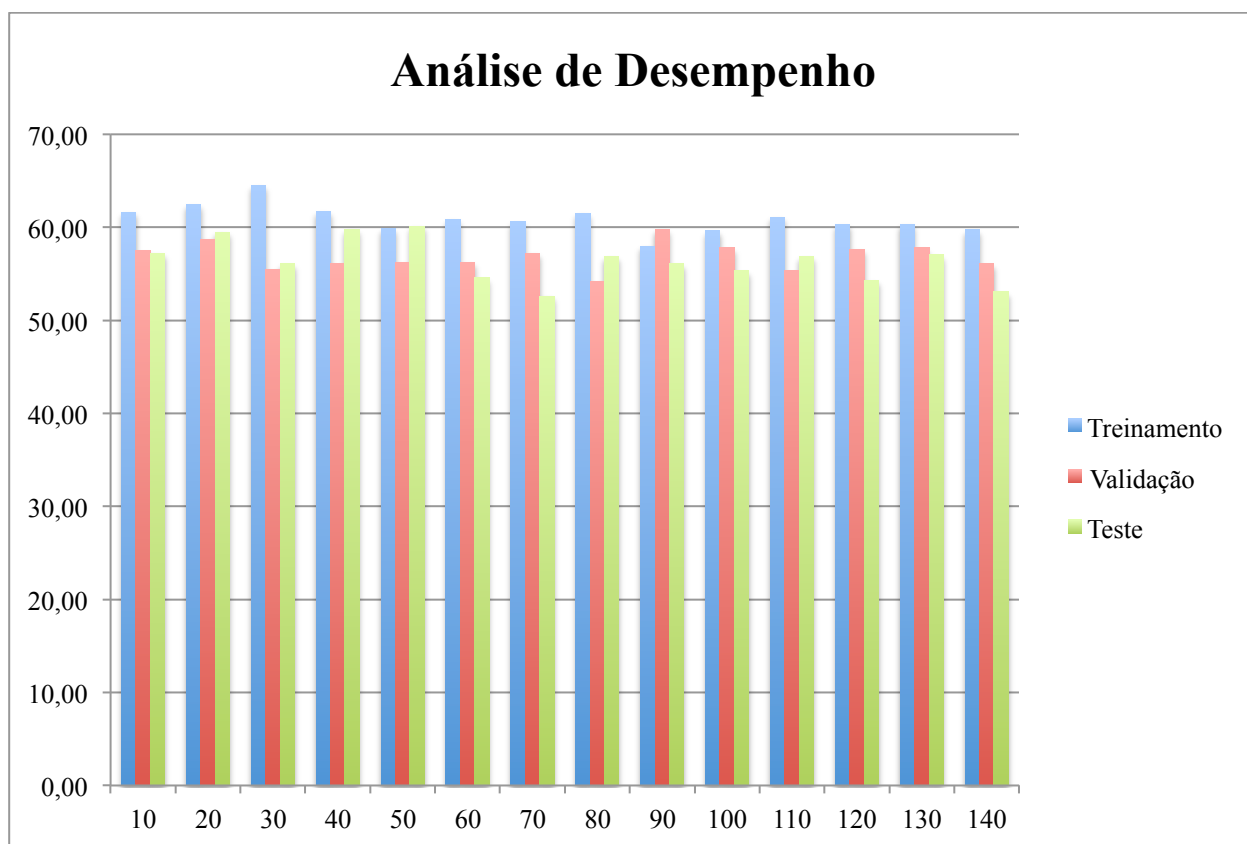


Figura 18: Gráfico de análise de desempenho

3.4 Outros experimentos

A princípio o estudo iniciou-se com menos variáveis na camada de entrada das RNAs, com apenas 4, como segue abaixo:

Variável	Primeira estudo	Estudo final
1	SKU (Identificação do produto)	SKU (Identificação do produto)
2	Semana	Semana
3	Quantidade de vendas na semana anterior	Quantidade de vendas na semana anterior
4	Preço médio semana	Preço médio semanal
5		Feriados comerciais / Período promocional
6		Tempo médio de entrega do pedido
7		Ações promocionais
8		Ações de Marketing

Tabela 23: Comparação entre quantidade de variáveis entre os estudos

Além da quantidade de variáveis, inicialmente os valores da base de teste eram valores decimais. Ao converter os valores decimais em valores binários, percebeu-se que a taxa de acertos aumentou. Logo, decidiu-se pela utilização de valores binários tanto para as variáveis de entrada quanto para os valores de saída esperados.

A quantidade de variáveis de entrada influenciou no resultado final dos estudos. O primeiro, com apenas 4 variáveis apresentou um resultado um pouco abaixo comparado ao estudo final que apresentou o dobro de variáveis na camada de entrada.

4 Conclusões

A pesquisa realizada teve o objetivo de elaborar um estudo aplicado utilizando redes neurais artificiais em sistemas de e-commerce para a previsão de vendas. O intuito desta pesquisa é mostrar que a inteligência artificial pode auxiliar gestores na tomada de decisão em um mercado competitivo e dinâmico.

As pesquisas realizadas de outros autores e os trabalhos correlatos mostram que ainda existe a necessidade de melhorar o desempenho dos resultados apresentados. A maioria dos trabalhos pesquisados tiveram dificuldade na normalização dos dados e problemas com os ruídos e inconsistências dos mesmos, dificultando o aprendizado e treinamento das redes neurais artificiais.

4.1 Resultados

Este trabalho mostrou que é possível elaborar redes neurais artificiais que façam a previsão de vendas futuras de produtos específicos de um e-commerce baseado em histórico de vendas.

A RNA que apresentou melhor resultado dentre as 14 redes neurais criadas, treinadas e testadas foi a rede com a arquitetura [43-50-4] apresentando 50 neurônios na camada intermediária. O resultado da rede com esta arquitetura foi de 60,09% de taxa de acertos. Isso significa que, se a área de vendas de um e-commerce selecionar um produto pertencente a base que foi utilizada na pesquisa e quiser saber a quantidade de vendas deste produto na próxima semana, a rede neural apresentará uma taxa de 60,09% de chance de acertar a previsão.

O resultado acima ainda está abaixo do desejado e a grande dificuldade são os fatores externos, como economia, período e promoções.

Uma dificuldade encontrada nesta pesquisa foi relacionada a base de dados. Eliminar ruídos da base de dados, selecionar variáveis da camada de entrada da rede neural, normalizar os dados foram tarefas onerosas nesta pesquisa.

4.2 Sugestões para Trabalhos Futuros

A pesquisa mostrou que é possível elaborar redes neurais que apresentem resultados satisfatórios. Uma sugestão para trabalhos futuros é seguir as mesmas metodologias apresentadas de obtenção de dados, seleção dos produtos, tratamento de dados, criação e análise, porém utilizando uma base de dados real de um e-commerce. Além disso, é interessante aumentar a quantidade de variáveis de entrada da rede neural. Nesta pesquisa foi necessário aumentar a quantidade de variáveis para resultar em um aumento na taxa de acertos.

Uma outra sugestão para trabalhos futuros é elaborar outras redes neurais com outros algoritmos de treinamento e comparar os resultados com as redes neurais multicamadas.

Referências Bibliográficas

BRAGA,A.P.; CARVALHO,A.P.L.F.; LUDERMIR,T.B. Redes Neurais Artificiais: Teoria e Aplicações. Editora LTC, 2000.

CHIANG,W.K; ZHOU, D.Z; ZHOU,L.Predicting and explaining patronage behavior toward web and traditional stores using neural networks: a comparative analysis with logistic regression, University of Maryland – Department of Information Systems, 2004.

CHOU,P.; LI,P.; CHEN,K.; WU, M. Integrating Web Mining and Neural Network for Personalized e-Commerce Automatic Service, Taiwan, 2009.

FERNEDA,E. Redes Neurais e sua aplicação em sistemas de recuperação de informação, Ci. Inf., Brasília, v. 35, n. 1, p. 25-30, jan./abr., 2006.

FREITAS,C.S; JÚNIOR,D.L.M; TANAKA,L.C; PIRES,D.F. Aplicação de Redes Neurais Artificiais em Sistemas de Apoio para Tomada de Decisão, Centro Universitário de Franca, Revista Eletrônica de Sistemas de Informação e Gestão Tecnológica, Vol. 01, Nr. 01, 2011.

GAMBOGI,J.A; COSTA,O.L. Aplicação de Redes Neurais na Tomada de Decisão no Mercado de Ações – IV Encontro de Administração da Informação, Bento Gonçalves, Rio Grande de Sul, 19 a 21 de maio de 2013.

HAYKIN,S. Redes Neurais Princípios e Prática. 2^a. ed. Editora: Bookman, 2001.

JUNIOR,A.N. Aplicação de Redes Neurais Utilizando o Software Matlab, Centro Universitário “Eurípedes de Marília”, Marília, 2005.

OLIVEIRA,T.C; JUNIOR,M.M.G. Sistema de Análise de Comportamento de Clientes Baseado em Rede Neural Artificial, Pontifícia Católica de Minas Gerais – Instituto de Ciências Exatas e Informática, 2012.

PASSARI,A. Exploração de Dados Atomizados para Previsão de Vendas no Varejo Utilizando Redes Neurais, Dissertação apresentada à Universidade de São Paulo – Departamento de Administração, São Paulo, 2003.

RUSSEL,S.J.; NORVIG,P. Inteligência Artificial (Terceira Edição). 3^a. ed. Editora: Campus, dezembro de 2009.

SANTOS,A.M; SEIXAS,J.M; PEREIRA,B.B; MEDRONHO,R.A. Usando Redes Neurais Artificiais e Regressão Logística na Predição da Hepatite A, Universidade Federal do Maranhão - Departamento de Matemática, Rev Bras Epidemiol, 2005.

Webgrafia

Oliveira, C. Disponível em: <<http://www.lem.ep.usp.br/Pef411/~Cristiano%20Oliveira/CristianoOliveira/Paginas/InteligenciaArtificial.htm>>. Acessado em 07 de dezembro de 2014.

Disponível em <http://www.decom.ufop.br/imobilis/wp-content/uploads/2012/06/03_Feedforward-e-Backpropagation1.pdf>. Acessado em 12 de março de 2015.

Simas, E. Processamento Estático de Sinais: Introdução as Redes Neurais Artificiais. Disponível em <<http://slideplayer.com.br/slide/292589/>>. Acessado em 08 de dezembro de 2015.

Disponível em <<http://www.cs.ru.nl/~ths/rt2/col/h10/10neurENG.html>>. Acessado em 08 de dezembro de 2015.

Attux, R. Disponível em <http://www.dca.fee.unicamp.br/~attux/topico_1.pdf>. Acessado em 08 de dezembro de 2015.

Watanabe, C.Y.V. Rede Perceptron: Capítulo 3. Disponível em <http://carolina.unir.br/downloads/1721_aula9_perceptron.pdf>. Acessado em 08 de dezembro de 2015.

Wang, H. Perceptron as a Principal Component Analyzer. Disponível em <<http://haohanw.blogspot.com.br/2014/12/ml-my-journal-from-neural-network-to.html>>. Acessado em 08 de dezembro de 2015.