

Aplicação de Redes Neurais na Previsão de Vendas para Retalho

INESC Porto

João Abel Vieira Sousa

Dissertação de Mestrado

Orientador na FEUP: Eng^o. Eduardo José Rego Gil Costa

Orientador no INESC Porto: Eng^o. Rui Diogo Rebelo



FEUP

Faculdade de Engenharia da Universidade do Porto

Mestrado Integrado em Engenharia Mecânica

2011-02-08

Aplicação de Redes Neurais na Previsão de Vendas para Retalho

Resumo

A variação das vendas ao longo do tempo é um dos maiores problemas que a indústria retalhista enfrenta hoje em dia. Para minimizar este problema as empresas tentam planejar as compras e distribuição dos produtos baseados na previsão dos padrões de vendas passados.

Nesta dissertação será feita uma tentativa de resolver este problema utilizando redes neuronais artificiais. O modelo de rede utilizado foi feedforward, e para tentar prever melhor as vendas utilizaram-se variáveis explicativas tais como, temperatura média, índice nacional de vendas de comércio a retalho, dia da semana, mês e saldos.

O trabalho efectuado nesta dissertação irá ajudar a melhorar a incerteza existente a curto/médio prazo no planeamento de vendas e distribuição.

Abstract

The variation of sales over time is a major problem that the retail industry faces nowadays. To minimize this problem companies try to forecast sales and distribution of products based on forecasts of sales patterns.

In this dissertation an attempt to resolve this issue will be made using artificial neural networks based on feedforward model, and to try to forecast sales variables such as average temperature , retail indices, day of the week, month and sales promotions were used.

This work will help to reduce the existing uncertainty in short and middle term planning of sales and distribution.

Agradecimentos

No decorrer deste trabalho foram muitas as pessoas que apoiaram, incentivaram e colaboraram na sua realização, às quais não posso deixar de manifestar o meu mais sincero agradecimento. Um primeiro agradecimento dirigido aos orientadores desta dissertação de mestrado:

- Ao Engenheiro Rui Diogo Rebelo, por toda a ajuda prestada e todo o apoio concedido, sempre disponível para o mais pequeno esclarecimento e para elevar o moral e motivação.
- Ao Engenheiro Eduardo José Rego Gil Costa, pelos conhecimentos que partilhou e pela ajuda na orientação desta dissertação.

À Doutora Patrícia Ramos e ao Engenheiro Pedro Souto por todo o apoio dispensado e à toda a equipa da UESP – INESC Porto.

Um agradecimento muito especial à minha família, à minha Mãe por todo o apoio nestes anos todos e à minha irmã que sempre esteve lá para mim. E finalmente obrigado Ana por todo o apoio e ajuda que me deste, um beijo.

Conteúdo

1.-Introdução.....	1
1.1.-Apresentação INESC Porto	1
1.1.1.-Apresentação da Unidade de engenharia e sistemas de produção (UESP).....	1
1.2.-Projecto de previsão de vendas.....	2
1.3.-Estrutura da Dissertação	2
2.-Estado da Arte	3
2.1. Introdução	3
2.2. Especificação da rede.....	5
2.2.1 . Arquitectura da rede	5
2.2.2. Processos de aprendizagem	7
2.2.3. Função de Activação	7
2.2.4. Medidas de eficiência da rede	8
2.2.5. Pré-tratamento dos dados	9
2.2.6. Amostra de treino e amostra de teste	10
2.2.7 Critérios de Paragem de Treino.....	10
2.3. Aplicação das redes neuronais artificiais na previsão.....	11
2.4.. Decomposição de séries temporais	11
2.4.1 Tendência	12
2.4.2. Sazonalidade.....	12
2.4.3. Irregularidades.....	12
2.5. Função de Autocorrelação Amostral (FAC).....	12
2.6 Função de Autocorrelação Parcial Amostral (FACP).....	13
3.-Descrição do problema.....	14
3.1.- Introdução.....	14
3.2. Dados utilizados na Dissertação	16
4.- Abordagem ao problema	17
4.1. -Introdução.....	17
4.2.- Definição do tipo de previsão.....	17
4.2 - Método para avaliação e comparação do desempenho dos modelos.....	22
4.2.1-Características da rede	24

Aplicação de Redes Neurais na Previsão de Vendas para Retalho

4.3.- Rede com input de vendas semanais e previsão semanal	25
4.3.1.- Definição dos <i>Inputs</i>	26
4.3.2. resultados.....	27
4.4. – Rede com <i>input</i> de vendas diário e previsão diária	27
4.4.1.- Variáveis explicativas	27
4.4.2.-Definição dos <i>inputs</i>	31
4.4.3.- Resultados	33
4.5.- Previsão de subcategorias	35
4.6.-Capacidade de generalização da rede	36
5.- Conclusão	39
5.1 - Conclusões.....	39
5.2.- Futuros Desenvolvimentos	40

Índice de figuras

Figura 1– técnica de janela deslizante nos inputs para capturar a sazonalidade	5
Figura 2- Uma rede típica do tipo <i>feedforward</i> (MLP).	6
Figura 3-Framework de previsão.....	15
Figura 4 - Agregação das vendas por estação, categoria e responsável	18
Figura 5 - Identificação das rupturas de stock.....	18
Figura 6 - Cálculo do peso da ruptura	19
Figura 7 - Estimação das vendas potenciais	19
Figura 8 -Evolução das rupturas acumuladas na estação Primavera/Verão	20
Figura 9 - Evolução diária das rupturas na estação Primavera/Verão 2009	20
Figura 10 – organigrama do sistema de distribuição.....	22
Figura 11 Ambiente de treino no programa Matlab R2010b®	23
Figura 12 Organigrama do processo de previsão utilizando RNAs	23
Figura 13 Exemplo de testes de arquitecturas para escolha de arquitectura ideal.....	24
Figura 14- Série temporal de vendas semanais de calçado da categoria sandália de uma loja na estação Primavera/Verão de 2007, 2008 e 2009.....	25
Figura 15 – Função de auto-correlação amostral e função de auto-correlação parcial amostral da série semanal.....	25
Figura 16 – técnica de janela deslizante nos inputs para capturar a sazonalidade	26
Figura 17- Série temporal de vendas diárias de calçado da categoria sandália de uma loja na estação Primavera/Verão de 2007, 2008 e 2009.....	28
Figura 18 – Função de auto-correlação amostral e função de auto-correlação parcial amostral da série diária.....	28
Figura 19 – Padrão de compra semanal.....	29
Figura 20 – Vendas diárias de calçado da categoria sandália de uma loja na estação Primavera/Verão de 2007, 2008 e 2009.vs temperatura média.	29
Figura 21 – Padrão mensal do índice de volume de negócios no comércio a retalho de têxteis, vestuário, calçado e artigos de couro entre Janeiro de 2007 e Dezembro de 2009 (Fonte: site do INE).	30
Figura 22 - Vendas diárias de calçado da categoria sandália de uma loja na estação Primavera/Verão de 2007, 2008 e 2009. Época normal e época de saldos	30
Figura 23 – estrutura da RNA utilizando os inputs propostos.....	32
Figura 24 – Previsão vs Realidade da loja que mais vende.....	33
Figura 25 EAM e REQM da amostra 1	35

Aplicação de Redes Neurais na Previsão de Vendas para Retalho

Figura 26 - Previsão vs Realidade de uma loja com correlação $0.6 >$ prevista utilizando a rede treinada da loja que mais vende.....	37
Figura 27- Previsão vs Realidade de uma loja com correlação > 0.7 prevista utilizando a rede treinada da loja que mais vende.....	38
Figura 28- Previsão vs Realidade de uma loja com correlação > 0.8 prevista utilizando a rede treinada da loja que mais vende.....	38

Índice de tabelas

Tabela 1 Funções de Activação	8
Tabela 2 Medidas de avaliação da exactidão da previsão.	9
Tabela 3 Fórmulas de normalização	10
Tabela 4- Motivos das rupturas	21
Tabela 5 – resultados da performance de previsão com horizonte temporal semanal.	27
Tabela 6 – resultados dos erros de previsão diárias utilizando variáveis explicativas.	33
Tabela 7- resultados das 3 amostras de 50 simulações cada	34
Tabela 8 – Previsão de subcategoria utilizando a rede 1	35
Tabela 9 – Análise de correlação das séries temporais de todas as lojas	36
Tabela 10 comparação da performance das previsões feitas com maior rede vs rede da própria loja	37

1.-Introdução

1.1.-Apresentação INESC Porto

O INESC Porto - Instituto de Engenharia de Sistemas e Computadores do Porto é uma associação privada sem fins lucrativos reconhecida como instituição de utilidade pública. Desenvolve actividades de investigação e desenvolvimento, consultoria, formação avançada e transferência de tecnologia nas áreas de Telecomunicações e Multimédia, Sistemas de Energia, Sistemas de Produção, Sistemas de Informação e Comunicação e Optoelectrónica.

O INESC Porto é uma instituição criada para constituir uma *interface* entre o mundo académico e o mundo empresarial da indústria e dos serviços, bem como a administração pública, no âmbito das Tecnologias de Informação, Telecomunicações e Electrónica, dedicando-se a actividades de investigação científica e desenvolvimento tecnológico, transferência de tecnologia, consultoria e formação avançada.

O INESC Porto procura pautar a sua acção por critérios de inovação, de internacionalização e de impacto no tecido económico e social, sobretudo pelo estabelecimento de um conjunto de parcerias estratégicas que garantam a sua estabilidade institucional e sustentabilidade económica.

1.1.1.-Apresentação da Unidade de engenharia e sistemas de produção (UESP)

A Unidade de Engenharia de Sistemas de Produção desenvolve a sua actividade nas áreas de Produção, (sistemas de informação avançados de apoio à gestão industrial; gestão da qualidade; gestão da manutenção; sistemas de planeamento e controlo da produção; racionalização e optimização dos processos produtivos; automação; sistemas de apoio à decisão) e Logística (sistemas de gestão de cadeias de fornecimento; planeamento de sistemas logísticos; integração e optimização de estruturas logísticas).

Dedica-se ainda ao Negócio Electrónico, Teletrabalho, Negócio Electrónico entre Empresas e Engenharia Empresarial (análise e optimização de processos no âmbito da gestão industrial; gestão de projectos de inovação empresarial; análise de requisitos técnicos e organizacionais; desenho e análise de redes de cooperação empresarial).

E tem os seguintes objectivos estratégicos:

- Contribuir para a melhoria do desempenho global de empresas industriais através da realização de projectos de I&D, consultoria, transferência de tecnologia e formação.
- Continuar a desenvolver competências em aspectos ligados aos sistemas de informação de apoio à gestão industrial, bem como aspectos relacionados com a organização das empresas.

1.2.-Projecto de previsão de vendas

O projecto consiste na concepção, desenvolvimento e implementação de um novo conceito organizacional orientado para a rede empresarial da cadeia de valor do calçado, que visa elevar para um patamar tecnológico inovador todos os procedimentos e sistemas de informação actualmente existentes, com base na integração de tecnologias RFID, que se estenderá ao nível :

- da fábrica de calçado;
- dos centros de armazenagem e redes de distribuição;
- das lojas comerciais e stands

O sistema utilizado no projecto com as características propostas terá como principais objectivos melhorar:

- a tendência ‘papel zero’ nos fluxos de informação de suporte a toda a rede;
- a rastreabilidade dos processos e produtos;
- armazenagem e distribuição ágil às lojas e entre lojas, adaptada à realidade das indústrias da moda designadamente, abertura das estações, reposições e logística inversa;
- a utilização de novos modelos de previsão que entrem em linha de conta com o cada vez maior número de novas variáveis que influenciam os hábitos de consumo;
- a tendência ‘stock zero’ ou seja, a redução dos stocks de produtos e materiais em toda a cadeia de valor, com conseqüente redução a montante de desperdícios materiais e mão de obra;
- a implementação de uma gestão operacional eficiente de logística inversa relativa aos ‘restos de colecção’ ;
- a captação de tendências e promoção dos produtos junto dos consumidores de forma mais dinâmica e interactiva.

1.3.-Estrutura da Dissertação

A presente dissertação é constituída por cinco capítulos, sendo este o capítulo introdutório ao problema que se pretender resolver.

O capítulo 2 é inteiramente dedicado ao estado da arte relativo às redes neuronais artificiais. O objectivo é apresentar de forma sucinta, mas clara, como funcionam as redes neuronais artificiais. Isto é realizado com o intuito de fornecer uma visão geral de como o problema da previsão pode ser resolvido.

No terceiro capítulo será descrito o problema que foi apresentado.

No capítulo 4 desta dissertação, através da análise do erro de previsão, será testado o desempenho da metodologia construída na resolução do problema.

As principais conclusões obtidas nesta dissertação são apresentadas no capítulo 5. Neste mesmo capítulo, são discutidas as dificuldades encontradas durante a realização desta dissertação e apresentadas sugestões para trabalhos futuros.

2.-Estado da Arte

2.1. Introdução

A previsão da procura no comércio de retalho é um dos aspectos fundamentais que está na base da tomada de decisões de estratégia e de planeamento de operações das organizações empresariais. Previsões de vendas precisas podem ajudar a tornar mais eficiente o funcionamento da cadeia de abastecimento, especialmente em grandes empresas com uma quota de mercado significativa. De uma má previsão poderá resultar um stock não desejável que afectará directamente o lucro e a posição competitiva da empresa (Zhang et al.,2005).

As vendas no comércio a retalho de calçado pertencem a um tipo especial de séries temporais que normalmente contêm ambos os padrões de tendência e sazonalidade, para além de outros aspectos, apresentando desafios para o desenvolvimento eficaz de modelos de previsão. A área científica da previsão tem sido dominada, desde à décadas, pelos métodos lineares. Os métodos lineares são relativamente fáceis de entender e, habitualmente, relativamente fáceis de implementar e interpretar. No entanto, têm uma capacidade muito limitada para capturar as relações não lineares dos dados. A utilização de modelos lineares para aproximar relações não lineares complexas, habitualmente presentes neste tipo de séries de dados reais, geralmente, não é satisfatória.

As redes neuronais artificiais (RNAs) têm sido utilizadas com sucesso para resolver muitos tipos de problemas de previsão (Zhang et al.,1998). O interesse crescente da sua utilização para a resolução de problemas práticos deve-se sobretudo à sua capacidade para capturar relações funcionais complexas, lineares ou não, subjacentes nas observações de séries temporais. Para além da capacidade versátil de modelação, as RNAs caracterizam-se também por serem uma metodologia não paramétrica “orientada aos dados”. Isto é, o procedimento em que se baseiam não exige pressupostos demasiado restritivos (estando menos susceptíveis a uma má especificação do modelo do problema que os métodos paramétricos) “aprendendo a partir dos dados ou experiência”. Esta metodologia de “aprendizagem” é muito interessante para problemas práticos onde habitualmente os dados são fáceis de obter e se conhece empiricamente que existem relações entre a variável a ser prevista e outros factores ou variáveis.

Nas últimas décadas têm sido desenvolvidos diversos modelos de RNAs para resolver diferentes tipos de problemas. As RNAs *perceptron* de múltiplas camadas (*multi-layer perceptrons-MLP* ou *feedforward-directas*) são sem dúvida as mais utilizadas para resolver problemas de previsão. Uma rede do tipo *MLP* é tipicamente composta por várias camadas de nós denominados neurónios artificiais. A camada de entrada (a mais baixa) é a camada onde a informação externa é recebida e a camada de saída (a mais alta) é a camada onde a solução do problema é obtida. As camadas de entrada e saída estão separadas por uma ou mais camadas intermédias denominadas camadas escondidas. Os nós das camadas adjacentes estão habitualmente ligados por arcos acíclicos orientados da camada mais baixa para a camada mais alta (Prasun Das et al.,2007).

Aplicação de Redes Neurais na Previsão de Vendas para Retalho

Para um problema de previsão causal as entradas da rede são habitualmente variáveis explicativas, podendo a relação funcional estimada pela rede ser escrita da forma:

$$y = f(x_1, x_2, \dots, x_p),$$

onde x_1, x_2, \dots, x_p são p variáveis explicativas independentes e y é a variável prevista dependente. Para um problema de previsão de séries temporais as entradas da rede são tipicamente as observações passadas da série de dados e o output a observação futura, podendo a relação funcional estimada pela rede ser escrita da forma:

$$y_{t+1} = f(y_t, y_{t-1}, \dots, y_{t-p}),$$

onde y_t é a observação no instante t . É também possível considerar como entradas da rede simultaneamente variáveis explicativas e observações passadas da série de dados. Uma RNA pode, assim, ser interpretada como um modelo auto-regressivo não linear para problemas de previsão de séries temporais.

Para que uma RNA possa ser usada para prever a variável desejada é necessário que seja treinada. Treinar uma rede neural consiste em determinar os pesos dos arcos e a polaridade dos nós de modo que, para uma dada entrada, a saída coincida com o valor desejado.

O treino para uma multilayer perceptron (*MLP*) é um treino supervisionado uma vez que o valor da saída (*target value*) correspondente a cada padrão de treino de entrada (*training pattern*). Para um problema de previsão causal o número de nós da camada de entrada é igual ao número de variáveis explicativas independentes.

Para um problema de previsão de séries temporais o número apropriado de nós da camada de entrada não é simples de determinar. Cada padrão de treino é sempre constituído por um “conjunto móvel de observações igualmente espaçadas” de comprimento fixo que percorre a série temporal.

Sendo, por exemplo, o conjunto de treino constituído por N observações y_1, y_2, \dots, y_N

e supondo que a previsão é a 1-passo, então, se a rede for composta por n nós de entrada:

- temos $N - n$ padrões de treino.
- O primeiro padrão de treino de entrada será constituído por y_1, y_2, \dots, y_n
- y_{n+1} será o nó de output (*target*);
- o segundo padrão de treino de entrada será constituído por y_2, y_3, \dots, y_{n+1}
- y_{n+2} será o nó de output;

e assim sucessivamente, finalmente:

- o último padrão de treino de entrada será constituído por $y_{N-n}, y_{N-n+1}, \dots, y_{N-1}$
- y_N será o nó de output.

Esta técnica designa-se de janela deslizante e tem por maior objectivo capturar a sazonalidade de uma série temporal.

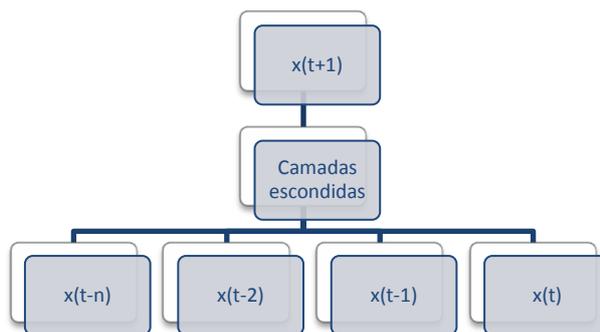


Figura 1– técnica de janela deslizante nos inputs para capturar a sazonalidade

O algoritmo de treino determina os pesos dos arcos da rede de modo a minimizar uma medida de erro global que pode ser, por exemplo, a média dos quadrados dos erros (MQE) ou a soma dos quadrados dos erros (SQE).

2.2. Especificação da rede

A especificação da arquitectura da RNA que melhor se adequa a um dado problema de previsão não é uma tarefa fácil. Essa especificação passa pela determinação do número de nós da camada de entrada, número de camadas intermédias e respectivo número de nós, e número de nós da camada de saída. Não existe nenhum método que determine directamente esses parâmetros, recorrendo-se habitualmente à simulação de um número limitado de cenários e a procedimentos do tipo tentativa-erro. Além destas especificações também temos de levar em conta outras opções de desenho da rede, função de activação, normalização dos dados, algoritmos de treino, tamanho dos *data-sets* de treino e de teste e medidas de performance da rede.

2.2.1 . Arquitectura da rede

Uma RNA é tipicamente composta por várias camadas de neurónios. A primeira é uma camada de input onde informação ou dados externos são recebidos, a última é uma camada de output onde obtemos a previsão. Estas camadas são separadas por uma ou mais camadas intermédias que são chamadas de camadas escondidas (*hidden layers*). Quando especificamos a arquitectura da rede temos que determinar os seguintes parâmetros:

- Numero de neurónios de *input*.
- Numero de camadas escondidas e neurónios escondidos
- Numero de neurónios de *output*

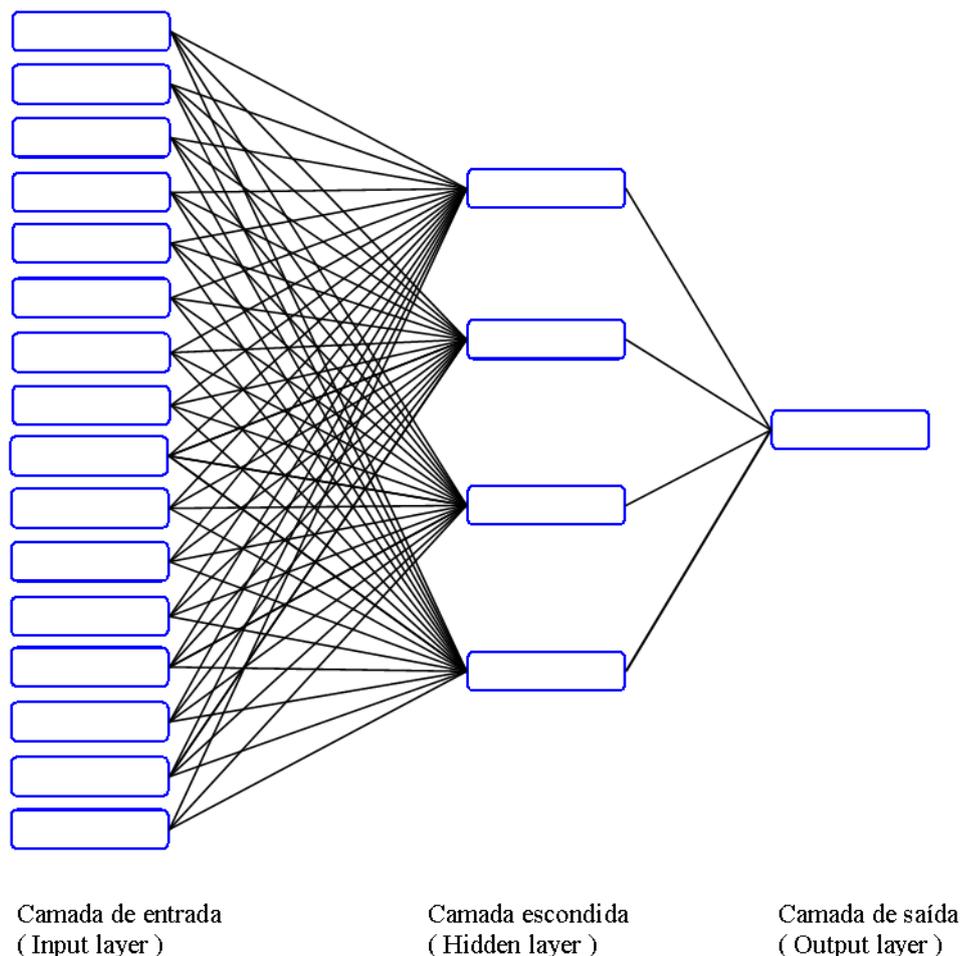


Figura 2- Uma rede típica do tipo *feedforward* (MLP).

A selecção destes parâmetros é dependente de cada problema.

Para um problema de previsão causal o número de nós da camada de entrada é igual ao número de variáveis explicativas independentes. O número apropriado de nós da camada de entrada não é simples de determinar uma vez que contém informação acerca da estrutura de auto correlação presente nos dados.

O número de camadas intermédias e respectivo número de nós da RNA é extremamente importante visto que são estes nós que permitem capturar o padrão de comportamento dos dados e determinar mapeamentos não lineares complexos entre as variáveis de input e as variáveis de output.

Resultados teóricos mostram que uma camada intermédia é suficiente para a rede determinar o mapeamento pretendido com a exactidão desejada. Contudo, pode ser necessário um número elevado de nós, o que poderá provocar um aumento do tempo de treino e poderá piorar a capacidade de generalização da rede.

Assim, para certo tipo de problemas, poderá ser benéfico utilizarem-se duas camadas intermédias. Há mesmo autores que simplesmente adoptam logo à partida duas camadas intermédias sem avaliarem a eficiência da rede apenas com uma camada.

De acordo com os estudos práticos já existentes e a conclusão de diversos autores (Lippmann (1987); Cybenko (1988); Chakraborty et al., 1992), uma RNA nunca necessita de mais do que duas camadas intermédias para resolver a maioria dos problemas, incluindo os de previsão.

Existem diversas heurísticas para determinar o número de nós intermédios.

Para o caso de RNAs apenas com uma camada intermédia, o número de nós intermédios é igual a

$$“2n+1”, “2n”, “n” e “n/2”,$$

sendo n o número de nós da camada de entrada.

No entanto nem todas estas heurísticas funcionam bem em todo o tipo de problemas.

O número de nós da camada de saída corresponde habitualmente ao horizonte da previsão. No caso de uma previsão a 1-passo somente é necessário um nó.

2.2.2. Processos de aprendizagem

Uma das características mais importante das RNAs é a capacidade de aprender com os dados que lhe são apresentados e com isso melhorar o seu desempenho. Esta operação é efectuada com o ajuste dos pesos das conexões de maneira a que se obtenha uma solução generalizada para um certo conjunto de dados e medida de eficiência da rede. O algoritmo de aprendizagem é de uma forma geral, um conjunto de regras que são seguidas para obter uma solução de optimização dos pesos da RNA.

O processo de aprendizagem, ou treino, das RNAs pode ser supervisionado ou não supervisionado. Na aprendizagem supervisionada a RNA aprende a partir dos dados históricos de entrada e dos dados correspondentes de saída, ou seja, a rede descobre uma relação entre os dados de entrada e os dados de saída.

O algoritmo mais utilizado para o treino supervisionado é designado por *backpropagation*, ou retropropagação do erro. De uma forma geral, este algoritmo compara a saída obtida pela rede neuronal com o valor da saída desejado, calculando o erro. Se o valor do erro for superior a uma determinada tolerância, é propagado para trás, alterando os pesos até se obter uma saída aceitável.

Na aprendizagem não supervisionada não existem exemplos para a aprendizagem de uma relação entre a entrada e saída, pois não são conhecidos os dados de saída. Este tipo de aprendizagem não é explorado nesta dissertação.

2.2.3. Função de Activação

A função de activação (também chamada função de transferência) determina a relação entre as entradas e as saídas de um nó e a rede. De um modo geral, a função de activação introduz um grau de não linearidade que é importante para a maioria das aplicações que usam RNAs. As funções de activação mais utilizadas são:

Tabela 1 Funções de Activação

Designação	Função	Contradomínio
Linear	$f(x) = x$	$]-\infty, +\infty[$
Sigmoid	$f(x) = \frac{1}{1+e^{-x}}$	$[0,1]$
Sigmoid1	$f(x) = \frac{2}{1+e^{-x}} - 1$	$[-1,1]$
Sigmoid2	$f(x) = \frac{x}{1+ x }$	$[-1,1]$
Tangente hiperbólica	$f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$	$[-1,1]$
Seno	$f(x) = \text{sen}(x) \quad (x \text{ mod } 2\pi)$	$[-1,1]$
Coseno	$f(x) = \text{cos}(x) \quad (x \text{ mod } 2\pi)$	$[-1,1]$
Gaussiana	$f(x) = e^{-\frac{x^2}{2}}$	$[-1,1]$

2.2.4. Medidas de eficiência da rede

A medida de eficiência da RNA mais importante é sem dúvida a exactidão da previsão. A exactidão da previsão é habitualmente definida em termos do erro de previsão que é diferença entre o valor efectivo e o valor previsto. As medidas habituais de avaliação da exactidão da previsão são:

Tabela 2 Medidas de avaliação da exactidão da previsão.

Designação:	
Erro absoluto médio (EAM):	$\frac{1}{n} \sum_{t=1}^n e_t $
Erro quadrático médio (EQM):	$\frac{1}{n} \sum_{t=1}^n (e_t)^2$
Raiz quadrada do erro quadrático médio (REQM):	$\sqrt{\frac{1}{n} \sum_{t=1}^n (e_t)^2}$
Erro percentual absoluto médio (EPAM):	$\frac{1}{n} \sum_{t=1}^n \left \frac{e_t}{y_t} \right \times 100$

e_t é o erro de previsão do instante t , y_t é o valor efectivo

N é o número de termos do erro

2.2.5. Pré-tratamento dos dados

Depois de escolher os dados mais apropriados estes têm de ser tratados, ou a RNA não irá ter uma performance adequada. As decisões feitas nesta fase de desenvolvimento são críticas para a performance da rede.

A transformação e a normalização são dois processos de pré-tratamento de utilização comum. A transformação envolve manipular os vários *inputs* da rede para criar um único input, enquanto que a normalização é uma transformação efectuada nos dados, para distribuí-los igualmente dentro de uma escala escolhida previamente.

Alguns métodos simples de pré-tratamento consistem em excluir dos *inputs* diferenças entre os mesmos e rácios, o que faz com que o número de *inputs* diminua e a rede aprenda mais rapidamente.

A normalização dos dados é o passo final no pré-tratamento. O objectivo é garantir que a distribuição dos valores para os inputs e outputs seja uniforme, também os valores devem ser normalizados de maneira a que o domínio dos dados inseridos na rede vá ao encontro do domínio da função de activação. As seguintes formulas são as mais frequentemente utilizadas para efectuar a normalização:

Tabela 3 Fórmulas de normalização

Transformação:	
Transformação linear para $[0, 1]$:	$x_n = \frac{x_0 - x_{\min}}{x_{\max} - x_{\min}}$
Transformação linear para $[a, b]$:	$x_n = (b - a) \frac{x_0 - x_{\min}}{x_{\max} - x_{\min}} + a$
Normalização estatística:	$x_n = \frac{x_0 - \bar{x}}{s}$
Normalização simples:	$x_n = \frac{x_0}{x_{\max}}$

x_n é o valor normalizado, x_0 é o valor original

$x_{\min}, x_{\max}, \bar{x}$ e s são o mínimo, máximo, média e desvio padrão do vector de dados original

2.2.6. Amostra de treino e amostra de teste

Na metodologia de RNAs os dados são normalmente subdivididos em dados de treino, validação e teste.

O conjunto de dados de treino é utilizado para ajustar os pesos das conexões, o conjunto de teste é utilizado para verificar os vários estágios de treino e determinar quando pará-lo para evitar over-fitting e o conjunto de validação é utilizado para estimar a performance de uma RNA treinada.

Segundo um estudo feito por Shahin et al (2004) onde foi investigado o impacto da proporção dos dados usados no treino, encontrou que não existia uma relação clara entre a proporção de dados para treino, teste e validação do modelo. No entanto, descobriu que os melhores resultados foram obtidos quando 20% dos dados totais eram utilizados para validação e os restantes eram divididos em 70% para treino e 30% para teste.

2.2.7 Critérios de Paragem de Treino

Os critérios de paragem determinam quando a RNA está optimamente treinada. Podem ser utilizados vários critérios para determinar quando parar o treino. O treino pode ser parado:

- após atingir um certo erro;
- quando não existe mudança de erro durante várias iterações;
- quando é treinado por um número fixo de padrões de treino.

No entanto, estes critérios de paragem podem levar à paragem prematura ou a um excesso de treino.

2.2.7.1 Excesso de Treino(*Over-fitting*)

Em redes em que o número de casos de treino não é adequado, caso a rede tenha sido submetida a um número excessivo de ciclos de treino, ocorre *over-fitting* (Geman, 1992), que consiste em a rede memorizar os padrões de treino e apresentar bons resultados para os casos de treino, embora para os casos de teste piore o desempenho, perdendo capacidade de generalização.

Este problema pode ser evitado escolhendo outras arquiteturas de rede, definindo outros critérios de paragem que façam com que a rede seja menos complexa e impeça de memorizar padrões.

2.3. Aplicação das redes neuronais artificiais na previsão

Uma das primeiras aplicações bem sucedidas utilizando RNA para previsão foi pelos autores Lapedes e Farber em 1987 e 1988, respectivamente. Estes mostraram que as RNAs podiam ser utilizadas para prever séries temporais não-lineares com muito boa eficácia. Depois destes autores muitos outros analisaram a capacidade das RNAs para prever séries temporais, com e sem ruído. As RNAs têm sido utilizadas muito frequentemente ao longo do tempo em previsões financeiras. Estas foram utilizadas para prever falências, variação das taxas de juros, preços de acções entre outros cenários financeiros.

Outra aplicação usual das RNAs é a previsão de consumo de carga eléctrica. Segundo os autores Park e Sandberg (1991), uma simples RNA só com *inputs* de temperatura tem uma performance melhor do que o método de regressão habitualmente utilizado.

Muitos outros problemas de previsão foram resolvidos utilizando RNAs, por exemplo:

- Propagação de pólen (Arizmendi et al., 1993);
- Preços de mercadorias (Kohzadi et al., 1996);
- Temperatura ambiental (Balestrino et al., 1998);
- Tráfego de passageiros numa companhia aérea internacional (Nam e Schaefer, 1995);
- Índices macroeconómicos (Maasoumi et al., 1994);
- Procura de água (Lubero, 1991).

2.4.. Decomposição de séries temporais

A decomposição de uma série temporal é utilizada para identificar vários padrões que simultaneamente a compõe. Estes padrões influenciam os dados e é muito importante estudar quais as influências que estes têm sobre a série temporal. De um modo geral existem três componentes na decomposição de séries temporais:

- Tendência (T_t);

- Sazonalidade (S_t);
- Irregularidades (I).

Os dois primeiros componentes são determinísticos (sinais), enquanto que a irregularidade é uma componente variável, também chamada de ruído.

Para ser feita uma previsão é importante saber quais destes componentes estão presentes nos dados.

$$X_t = S_t \cdot T_t \cdot C_t.$$

onde X_t representa o valor da série temporal.

2.4.1 Tendência

A tendência é o crescimento ou o decréscimo de uma série temporal a um ritmo constante ao longo do tempo. As séries temporais que apresentem este tipo de componente são chamadas de não estacionárias. Este comportamento ocorre quando várias situações se conjugam, como por exemplo, inflação, crescimento da população ou aumento de produtividade.

2.4.2. Sazonalidade

A sazonalidade é uma flutuação que é repetida anualmente, aproximadamente na mesma altura e com o mesmo nível de intensidade. Esta componente é frequente em produtos que só se vendem em certas alturas do ano ou que são muito dependentes das condições meteorológicas.

2.4.3. Irregularidades

Esta componente das séries temporais é imprevisível e é uma das fontes de erro ao fazer previsão de séries temporais.

2.5. Função de Autocorrelação Amostral (FAC)

A Função de Autocorrelação Amostral é definida como a correlação que a série temporal tem com ela própria, deslocada de k períodos de tempo.

$$\hat{p}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=1}^{n-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{\sum_{t=1}^n (Z_t - \bar{Z})^2}, k = 0, 1, 2, \dots, n$$

$$\text{onde } \bar{Z} = \sum_{t=1}^n Z_t / n.$$

A FAC permite perceber a dependência estatística entre os dados.

2.6 Função de Autocorrelação Parcial Amostral (FACP)

A Função de Autocorrelação Parcial Amostral é a autocorrelação entre a observação actual no tempo t e a mesma desfasada de $t+k$ periodos, assim:

$$\hat{\phi}_{k+1,k+1} = \frac{\hat{p}_{k+1} - \sum_{j=1}^k \hat{\phi}_{kj} \hat{p}_{k+1-j}}{1 - \sum_{j=1}^k \hat{\phi}_{kj} \hat{p}_j}$$

$$\text{e } \hat{\phi}_{k+1,j} = \hat{\phi}_{kj} - \hat{\phi}_{k+1,k+1} \hat{\phi}_{k,k+1-j}, j = 1, \dots, k$$

3.-Descrição do problema

3.1.- Introdução

As redes de retalho moda/calçado são tipicamente caracterizadas por um *backoffice*, uma unidade central de armazenamento, um sistema logístico de distribuição e uma rede de lojas. A unidade central tem um papel fundamental, assumindo responsabilidade, entre outras, na definição das quantidades de produtos a enviar para cada loja, nas transferências de artigos entre lojas e na elaboração das encomendas.

É neste tipo de redes de retalho que este projecto se pretende posicionar, propondo-se a resolver as fragilidades que actualmente existem nestes modelos.

Actualmente verifica-se que as práticas de previsão de vendas, nas pequenas e médias empresas comerciais, incluindo cadeias de lojas de retalho de produtos de moda, mantêm-se geralmente pouco ambiciosas, pouco elaboradas, pouco inovadoras, e também pouco eficazes.

As previsões estão ainda muito suportadas nas questões inerentes à oferta (*supply elements*), minimizando a importância da procura (*demand elements*) num panorama onde os consumidores são cada vez mais exigentes e conhecedores das suas alternativas.

Apesar de estar muito generalizada a adopção de práticas de fidelização de clientes, tais práticas não estão associadas a metodologias globalizantes que permitam segmentações apropriadas da clientela, ou que perspectivem o registo apropriado do que os clientes compram, preferem ou gostavam de adquirir..

Não existem modelos de previsão minimamente apropriados à realidade, fáceis de utilizar, flexíveis, adaptativos, que reconheçam a necessidade de fazer chegar às lojas, permanentemente, os produtos que melhor servem os anseios dos consumidores.

Os dados necessários para se efectuar uma boa previsão de vendas estão disponíveis, mas poucas são as empresas que o fazem de maneira estruturada. O emprego de algoritmos de redes neurais artificiais tem-se mostrado uma opção promissora no sentido de converter os dados disponíveis numa previsão de vendas útil (Kin-Fan Au, 2008).

Apesar das vantagens apresentadas na literatura sobre redes neurais, Armstrong (2001) salienta que não existe nenhum método de previsão que se adequa a todas as situações, pelo que é importante saber estruturar e representar realisticamente o problema antes de testar a eficácia dos modelos.

A falta de modelos eficientes de previsão da procura é geradora de importantes desperdícios a diversos níveis, nomeadamente referências excessivas, sortidos mal planeados, sobras excessivas de colecções, imprópria disponibilidade nas prateleiras, perdas de vendas, baixas rotações de stock, deficiente introdução de novos produtos, falta de planeamento atempado de promoções e campanhas, entre outros. O objectivo desta dissertação centra-se na concepção e desenvolvimento de uma nova plataforma de previsão para o sector do calçado utilizando RNAs. Será criada uma estrutura de recolha de dados qualitativos que serão alvo de tratamento estatístico e concebida uma solução de previsão inovadora que se adequa às características e especificidades da previsão da procura para o sector Moda Calçado.

Aplicação de Redes Neurais na Previsão de Vendas para Retalho

Pretende-se assim aumentar a eficácia e o grau de automatização no processo de decisão e uma nova metodologia de monitorização dos factores críticos de venda.

A informação resultante deste sistema será de fulcral importância para qualquer rede de distribuição, pois servirá de base ao desenvolvimento do produto, à negociação com os fabricantes, ao planeamento do negócio e à rede de distribuição das lojas.

A plataforma de previsão a desenvolver requererá investigação a dois níveis: modelos matemáticos e estruturação de informação quantitativa e qualitativa.

No âmbito dos modelos matemáticos será efectuada uma investigação essencialmente ao nível das RNA e a sua aplicabilidade no problema definido.

Ao nível da estruturação da informação serão consideradas informações:

- enquadramento (visibilidade, localização, áreas de exposição, factores climatéricos, condicionantes geográficas, legais, eventos de proximidade)
- gestão de produtos (coleções, categorias, estratégias, táticas de promoção, área de exposição)
- clientes (Factores demográficos, geográficos, psicológicos, comportamentais, etários, assiduidade, poder de compra, etc);

Estes dados devidamente tratados e analisados permitirão criar modelos de previsão com diferentes níveis de agregação e diferentes horizontes temporais e melhorar o controlo dos factores críticos da empresa.

Em conclusão pretende-se a implementação de um sistema que permitirá gerir de uma forma eficiente os aprovisionamentos, a produção, as redes de distribuição e as lojas de venda, sobretudo os pertencentes aos segmentos moda, fazendo face às crescentes exigências dos mercados e consumidores .



Figura 3-Framework de previsão

3.2. Dados utilizados na Dissertação

Os dados utilizados nesta dissertação foram fornecidos por uma empresa líder na indústria de calçado em Portugal, que tem uma vasta rede de vendas ao retalho, bem distribuída pelo país, com mais de 50 lojas. Por questões de confidencialidade os dados foram alterados, mas representam a realidade e não influenciam a demonstração do resultado.

Foi fornecida, pela empresa, uma base de dados contendo informação relativa a vendas, produtos, categorias, lojas, saldos, armazéns e stocks.

Toda a informação fornecida pela empresa foi tratada utilizando o programa Microsoft SQL Server 2008 R8.

Para testar se as RNA têm aplicabilidade na previsão de vendas na indústria do calçado, foram estudadas séries temporais de vendas de categorias.

Ao nível de dados obteve-se:

- Histórico de entradas diárias de artigos em armazém;
- Histórico de vendas diárias ocorridas para cada artigo em cada loja;
- Histórico de movimentações diárias ocorridas para cada artigo em cada loja;
- Ficha descritiva de cada artigo.

4.- Abordagem ao problema

4.1. -Introdução

Neste trabalho foi desenvolvido um modelo para previsão das vendas diárias de calçado de uma empresa nacional utilizando redes neurais artificiais.

Numa primeira fase foram analisados os perfis de comportamento de diversas séries de vendas de calçado de diversas categorias em diferentes lojas.

Com o intuito de desenvolver um modelo de previsão causal foram inspeccionadas as relações entre as quantidades vendidas diariamente e algumas variáveis consideradas “explicativas”, tais como

- temperatura média,
- precipitação total,
- índice de volume de negócios no comércio a retalho de têxteis, vestuário, calçado e artigos de couro, e
- disponibilidade de artigos/modelos em loja, dia da semana, promoções.

A implementação da rede neuronal artificial foi levada a cabo usando o software *Matlab R2010b*®.

Após a análise das séries temporais das vendas, será feita uma análise das rupturas de stock para tentar avaliar o horizonte temporal de previsão. Feito este estudo o passo seguinte será testar várias arquitecturas e vários inputs e tentar perceber quais terão melhores resultados.

4.2.- Definição do tipo de previsão

Inicialmente procurou-se definir a complexidade do problema associado à previsão, e foi necessário perceber qual o melhor horizonte temporal mais adequado:

- um horizonte longo para fazer a previsão ao nível das compras, ou
- um horizonte curto para melhorar a distribuição.

Com esse objectivo, foram estudadas as rupturas de stock, pois as vendas de um artigo estão limitadas pela quantidade disponível em stock, o que em caso de ruptura não traduz a procura efectiva do artigo.

Foram assumidos os seguintes pressupostos no estudo das rupturas de stock:

- a ruptura de stock de um artigo numa loja acontece quando o stock é nulo para todos os tamanhos desse artigo, na referida localização
- o fim de cada estação corresponde à data final máxima para um período de ruptura.

A análise foi restringida a uma marca, ao mercado senhora e ao horizonte temporal de 27/02/2007 a 27/02/2010.

Calculou-se o stock final após cada movimento de entrada, saída ou venda em cada loja para cada artigo.

$$\text{Stock Final} = \text{Stock Inicial} + \text{Entradas} - \text{Vendas} - \text{Saídas}$$

Em seguida foram agrupadas as vendas diárias por categoria e responsável em cada estação (Figura 4).

Estação	Data	Artigo	Loja	Vendas			
2007V	07-06-2007	a	col	3	Categoria: Sapatos	Responsável: Gina	
2007V	07-06-2007	b	col	1		Categoria: Sandália	Responsável: Paulo
2007V	07-06-2007	b	csc	2			
2007V	07-06-2007	b	opk	1			
2007V	07-06-2007	b	nor	4			
2007V	07-06-2007	c	nor	1			

Figura 4 - Agregação das vendas por estação, categoria e responsável

Esta agregação deveu-se ao facto da movimentação de artigos entre lojas ser efectuada, na generalidade dos casos, pelos responsáveis e de as vendas de certos artigos nas lojas serem residuais, o que dificulta ter uma visão correcta do comportamento de vendas.

A etapa seguinte correspondeu à identificação das rupturas de stock, isto é, de acordo com o pressuposto assumido, quando um artigo em todos os tamanhos não está disponível para venda em loja (Figura 5).

Uma ruptura inicia-se com todas as movimentações de venda ou saída (excepto para armazém) que originaram stock zero. O fim de uma ruptura ocorre quando há uma movimentação de entrada que coloca o stock positivo ou então é considerado o último dia da estação.

Estação	Artigo	Categoria	Loja	Responsável	Data Inicio Ruptura	Data Fim Ruptura	Dias de ruptura
2007V	a	Sapatos	col	Gina	07-04-2007	10-04-2007	3
2007V	a	Sapatos	col	Gina	15-04-2007	20-04-2008	5
2007V	b	Sapatos	opk	Paulo	07-06-2007	21-06-2007	14
2007V	c	Sandália	col	Gina	05-04-2007	12-04-2007	7

Figura 5 - Identificação das rupturas de stock

Em seguida foi calculado o peso da ruptura, que indica a razão entre vendas ocorridas no período de ruptura e as vendas totais na estação, para a mesma categoria a que o artigo pertence e para o mesmo responsável da loja (Figura 6). Assim identificou-se o comportamento esperado de artigos semelhantes em lojas geridas pelo mesmo responsável.

Aplicação de Redes Neurais na Previsão de Vendas para Retalho

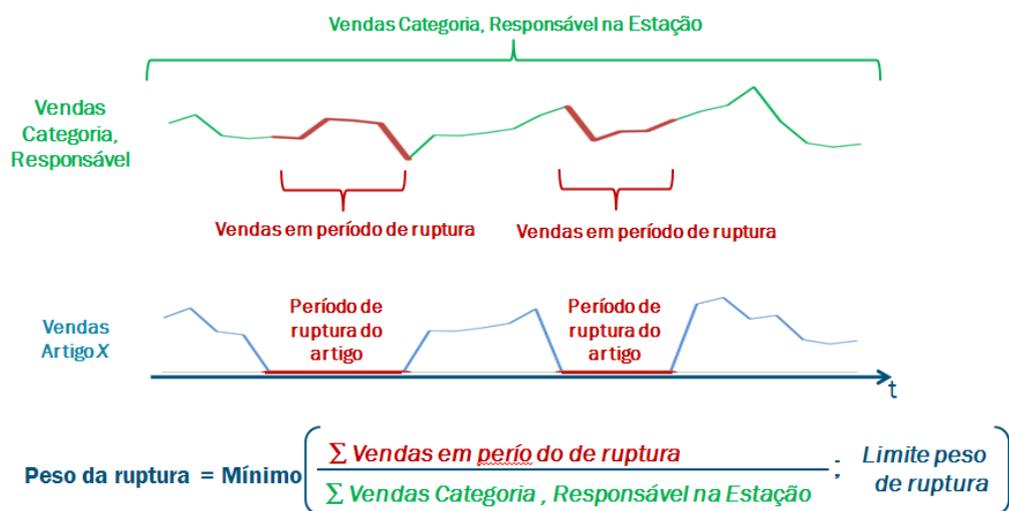


Figura 6 - Cálculo do peso da ruptura

Assumiu-se que deveria existir um limite para o peso de ruptura verificado, pois os artigos não repetem o mesmo perfil de venda ao longo da estação, assim como se considerou que os clientes também poderiam optar por artigos com características semelhantes face à inexistência do produto pretendido.

Por último foram estimadas as vendas potenciais de um artigo em determinada loja caso não existissem rupturas (Figura 7). As vendas potenciais correspondem à razão entre as vendas reais ocorridas e o somatório do peso das vendas fora do período de ruptura.



Figura 7 - Estimação das vendas potenciais

Após o cálculo dos indicadores de rupturas e das vendas potenciais procurou-se analisar graficamente os resultados.

A Figura 8 apresenta a evolução por dia de estação das rupturas acumuladas de artigos em lojas, com distinção entre época normal de vendas e de saldos. Observa-se que 50% das rupturas ocorreram na época normal de vendas quando se esperaria um maior número de ocorrências no período de saldos.

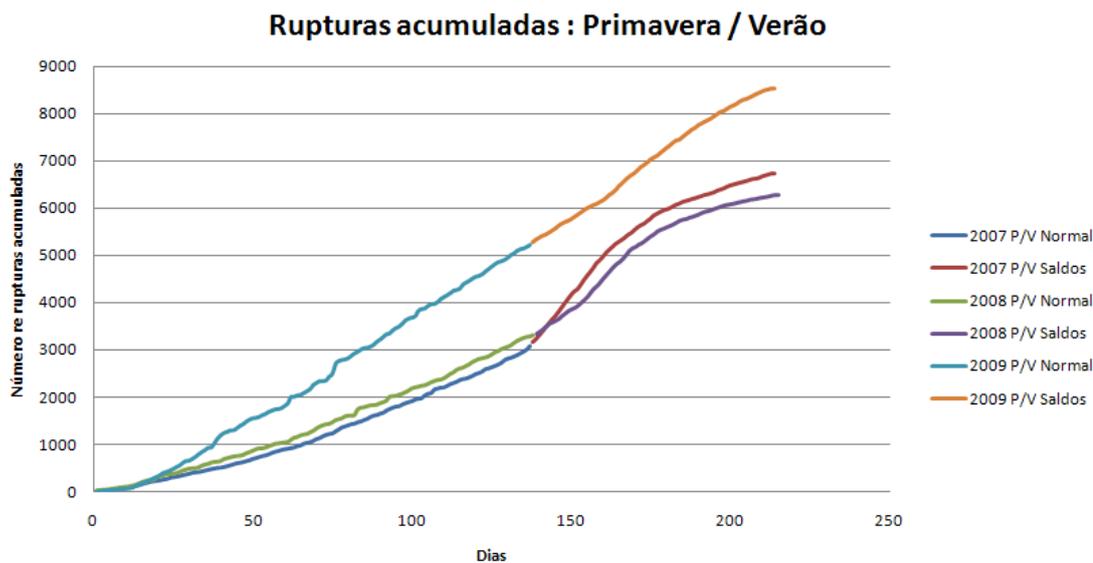


Figura 8 -Evolução das rupturas acumuladas na estação Primavera/Verão

A Figura 9 mostra, usando como exemplo a estação Primavera/Verão 2009, a evolução diária das rupturas. A vermelho está representado o número de rupturas que foram corrigidas (que deixaram de existir), enquanto que nos tons de azul claro e azul escuro se representa o número de artigos em ruptura em cada dia consoante a época da estação (normal ou saldos).

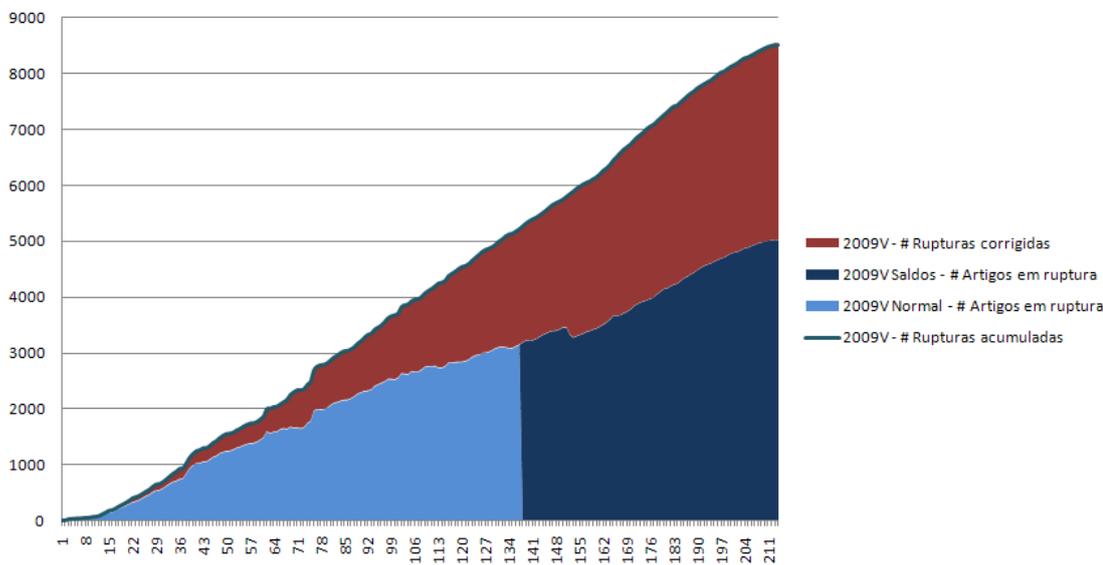


Figura 9 - Evolução diária das rupturas na estação Primavera/Verão 2009

Aplicação de Redes Neurais na Previsão de Vendas para Retalho

Verifica-se que no final da época normal de vendas, das 5000 rupturas acumuladas, apenas 2000 foram corrigidas e 3000 mantêm-se por resolver. Considera-se que a compra é insuficiente quando a quantidade encomendada e recebida de um artigo é inferior às vendas potenciais, sendo o oposto considerado uma distribuição não eficaz.

A Tabela 4 mostra que a distribuição é a grande responsável da maior parte das rupturas.

Tabela 4- Motivos das rupturas

Estação	Compra insuficiente	Distribuição não eficaz
2007 Outono/Inverno	5 %	95 %
2008 Outono/Inverno	4 %	96 %
2009 Outono/Inverno	3 %	97 %
2007 Primavera/Verão	16 %	84 %
2008V Primavera/Verão	9 %	91 %
2009V Primavera/Verão	15 %	85 %

Através desta análise identificou-se que a maior parte da origem das rupturas está associada a problemas de distribuição, ou seja, o nosso estudo será direccionado para a melhoria da previsão de vendas em cada loja.

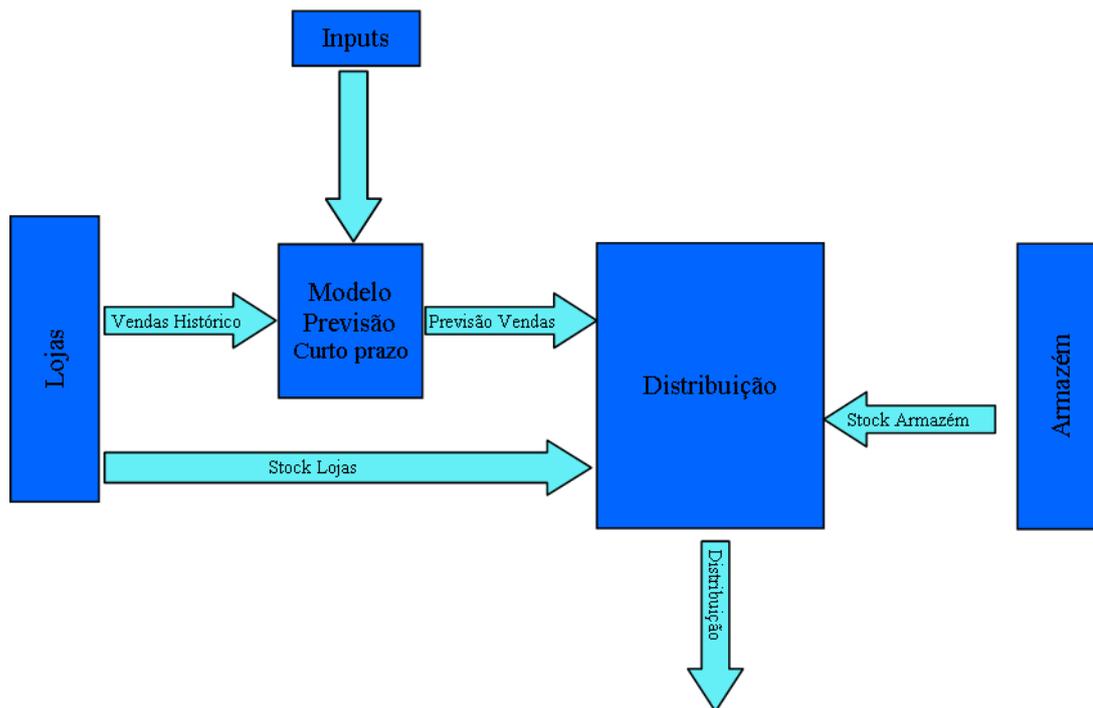


Figura 10 – organograma do sistema de distribuição

4.2 - Método para avaliação e comparação do desempenho dos modelos

Um dos factores mais importantes ao construir uma RNA é tomar a decisão sobre o que a rede vai “aprender”. O grande desafio será determinar quais os *inputs* que serão utilizados e reunir dados suficientes para treinar a rede convenientemente. Os dados de *input* devem permitir que a RNA generalize o comportamento do mercado.

Para verificar se a utilização de RNAs é viável para fazer previsão de vendas no calçado será testada a capacidade de performance da rede treinada com dados que a rede desconhece, ou seja, dados que não foram utilizados para o treino.

O treino será feito com dados relativos às vendas da categoria mais relevante para a empresa no período de análise, Primavera/Verão 2007, 2008, 2009, onde o período de 2007 a 2008 será utilizado para treino e o período de 2009 para validação da RNA.

Posteriormente será feita uma análise do EAM, do EQM e do REQm.

As RNAs das séries temporais de vendas diárias de calçado das diferentes categorias foram treinadas e simuladas utilizando o programa Matlab R2010b®.

Aplicação de Redes Neurais na Previsão de Vendas para Retalho

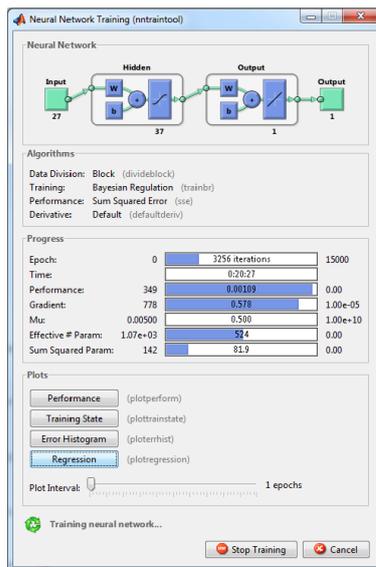


Figura 11 Ambiente de treino no programa Matlab R2010b®

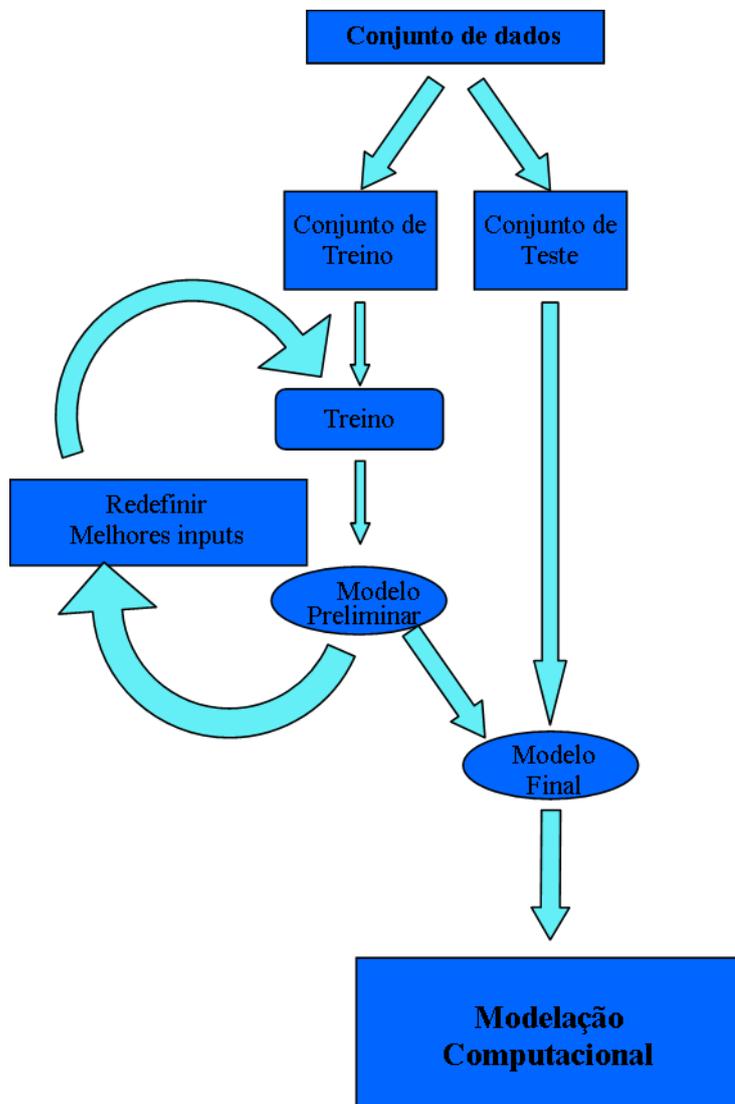


Figura 12 Organigrama do processo de previsão utilizando RNAs

4.2.1- Características da rede

Para definir a rede a ser utilizada, tivemos que escolher o algoritmo de treino, número de layers, neurónios de cada layer, função de transferência, tipo de normalização dos dados e divisão dos dados para treino e validação.

Relativamente ao algoritmo de treino, na fase inicial dos testes foi utilizado o algoritmo *Levenberg-Marquardt (trainlm)*, mas mais tarde após comparecer a uma apresentação, na Faculdade de Ciências da Universidade do Porto, de um trabalho intitulado “modelo para previsão do consumo diário de gás natural” feito pelo Doutor Nuno Carmona onde o algoritmo explorado era o *Bayesian Regulation Backpropagation (trainbr)*. Apesar de ser mais lento, o algoritmo *trainbr* apresentou melhores resultados, sendo portanto o escolhido.

Alguns autores como Caudil (1990), apresentam o número de neurónios das *hidden layers* como incógnita, e explicam que o processo para a sua escolha é um processo de tentativa-erro, pois se tivermos muitos neurónios a rede estará a memorizar os padrões e a perder capacidade de generalização (*over-fitting*). Por outro lado, se tivermos poucos neurónios, a rede demorará mais iterações a atingir o erro desejado.

No entanto, manter o número de neurónios baixo desde que providencie resultados satisfatórios é sempre melhor, dado que reduz o tempo computacional necessário para o treino, ajuda a RNA a obter melhor generalização e ajuda a evitar o problema do *over-fitting*.

Relativamente ao número de *layers*, embora com uma *layer* fosse possível fazer uma previsão adequada, de tempo demorado para cada simulação, optou-se por utilizar duas *layers* que também trazem vantagens quando tratamos de problemas mais complexos. O número de neurónios por *layer* foi definido utilizando um programa que testava as várias arquiteturas admissíveis com diferentes configurações.

Foram feitas 205 simulações com arquiteturas diferentes com 2 *layers* e conseqüente análise de erros para cada rede, chegou-se a conclusão que a rede ideal para duas *layers* era constituída por 4 neurónios na primeira *hidden layer* e 3 neurónios na segunda *hidden layer*.

	EQM	EAM	REQM	Arquitecturas
	5,369133	4,321478	20,34927	5321
	6,19036	5,183223	24,6678	515111
	4,846948	3,887031	29,1926	515121
	5,954445	4,879785	31,34966	515131
	4,715431	3,700693	24,15397	515141
	8,274491	5,740142	32,03105	515151
	3,800277	3,107	15,52679	Erros mínimos
Arquitecturas Ideais	5-14-8-1	5-11-5-1	5-15-5-1	

Figura 13 Exemplo de testes de arquiteturas para escolha de arquitetura ideal

A função de transferência utilizada foi a função *sigmoid*. Os dados foram normalizados para o domínio [0,1].

4.3.- Rede com input de vendas semanais e previsão semanal

Para um estudo inicial começou-se por analisar as vendas da categoria mais vendida, na loja que apresentava maiores vendas da empresa. Os valores negativos, que significavam que houve ruptura de stock naquele dia, foram substituídos por 0. Todos os dias em que não existiam vendas foram também substituídos por 0.

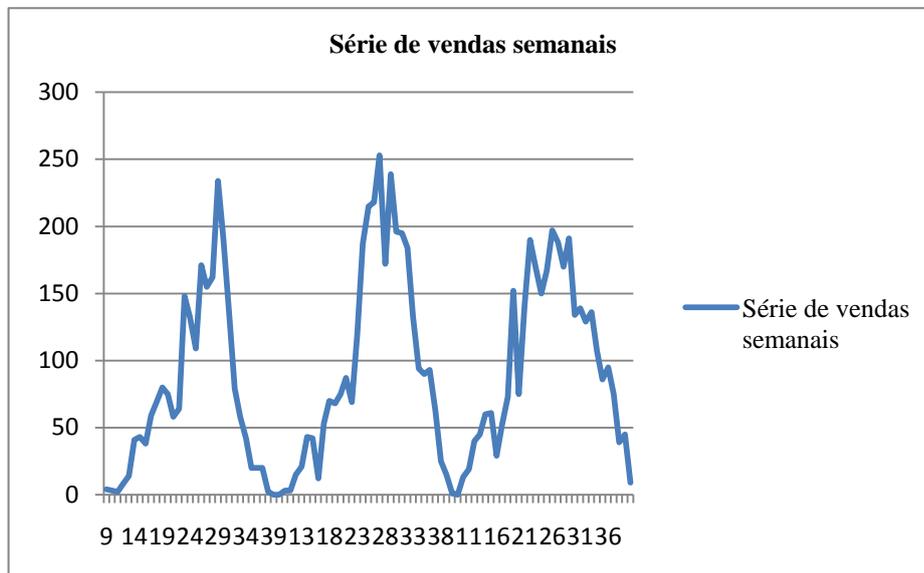


Figura 14- Série temporal de vendas semanais de calçado da categoria sandália de uma loja na estação Primavera/Verão de 2007, 2008 e 2009

As vendas foram agrupadas semanalmente, num total de 99 semanas no período de estudo, sendo 66 utilizadas para dados de treino e as restantes 33 para validação da previsão efectuada.

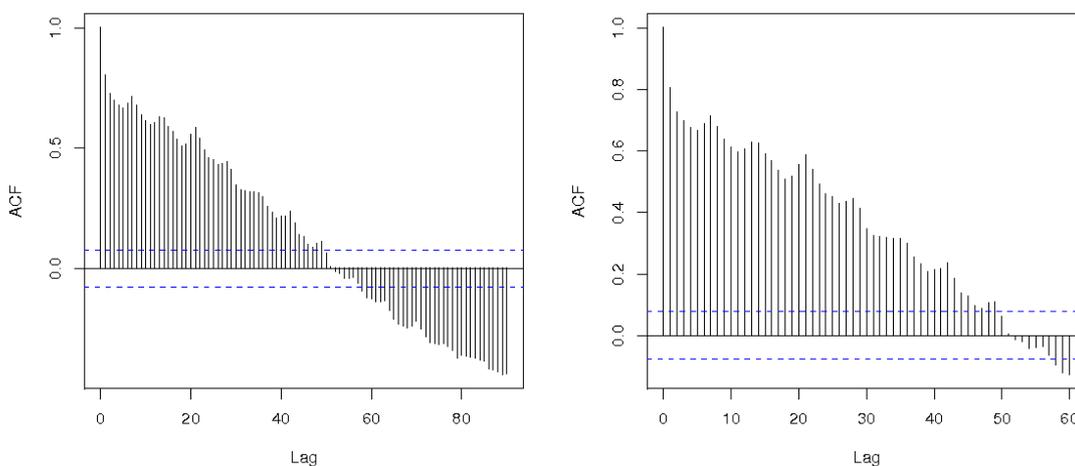


Figura 15 – Função de auto-correlação amostral e função de auto-correlação parcial amostral da série semanal

O gráfico da respectiva função de auto-correlação amostral (Figura 15) mostra que:

- a série é não estacionária (os valores da auto-correlação mantêm-se elevados e positivos);
- a série é sazonal (os valores da auto-correlação para os *lags* 7, 14, 21,28 são sempre superiores às auto-correlações adjacentes) sendo a sazonalidade semanal.

Como a auto-correlação amostral sugere, existe sazonalidade semanal, o que nos vai levar a que os primeiros testes de previsão sejam feitos com um horizonte temporal semanal.

4.3.1.- Definição dos *Inputs*

Quando tentamos replicar um comportamento sazonal de uma série numa RNA, é crítico incluir nos *inputs* as observações separadas por períodos.

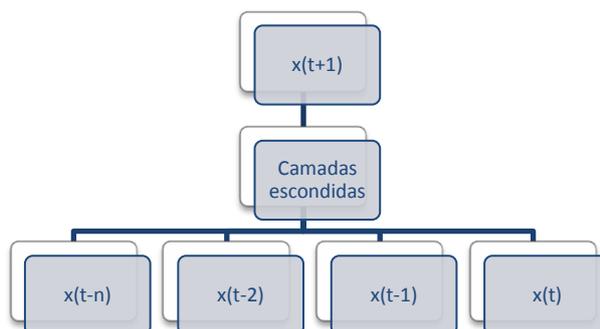


Figura 16 – técnica de janela deslizante nos inputs para capturar a sazonalidade

Para esta análise inicial foi testada a capacidade preditiva da RNA utilizando as vendas de 4 semanas consecutivas como *input* para prever as vendas da semana seguinte e analisando o EAM, o EQM e o REQM. Foi também testada a hipótese de 5 semanas e 6 semanas como janela temporal. Posteriormente foram adicionados a temperatura média semanal e a pluviosidade média semanal para tentar melhorar a performance da RNA.

Utilizando a janela de 4 semanas o conjunto de dados, que inicialmente era um vector das vendas 1×159 , foi normalizado e depois transformado numa matriz de 4×155 . Posteriormente foi dividido em duas matrizes, uma de treino e outra para validação dos testes, sendo a matriz de *inputs* para treino definida em 4×102 , enquanto que o vector de output foi definido por 1×102 .

Depois da RNA treinada, foi feita a simulação com as restantes 53 semanas, e medido a performance da previsão com base na análise dos erros.

Este processo foi efectuado analogamente para a janela de 5 e 6 semanas respectivamente. Para incluir a temperatura média semanal foi adicionado mais uma coluna de *input* com a respectiva variável.

4.3.2. resultados

Analisando a tabela 5 podemos verificar que a previsão utilizando a Temperatura média semanal apresenta melhor resultado quer a nível de EAM quer REQM. No entanto dado que a previsão semanal apresenta um erro significativo, e como o horizonte de previsão é grande, tendo em conta que esta previsão destina-se a alimentar os algoritmos de optimização de rotas de distribuição à loja e logística inversa, optou-se por abandonar a previsão com horizonte semanal e iniciar testes com horizonte de previsão diário.

Tabela 5 – resultados da performance de previsão com horizonte temporal semanal.

Inputs				
só vendas	Arquitectura da Rede	EQM	REQM	EAM
4 semanas	4421	402.4799	20,0619	14,8391
5 semanas	51421	442.4376	21,0342	15,9036
6 semanas	6651	398.2483	19,95616	15,752747
Temperatura média semanal				
4 semanas	5661	329.7180	18,158137	13,631141
5 semanas	6221	413.1670	20,32651	15,50415
6 semanas	7731	440.8059	20,995378	15,694375

4.4. – Rede com *input* de vendas diário e previsão diária

Para que a RNA capture e aprenda as características da série temporal, a quantidade de dados disponível deve ter uma grande variedade, de forma à RNA ter uma boa capacidade para generalizar. Com isto em mente, testamos a capacidade da RNA fazer previsão das vendas da categoria mais vendida diariamente. Assim passou-se a ter 1095 padrões de treino equivalentes aos dias entre 01/01/2007 a 31/12/2009 o que aumenta consideravelmente o número disponível de dados e por conseguinte, a capacidade de aprendizagem da rede.

4.4.1.- Variáveis explicativas

Analisando a série temporal da sandália na loja que mais vende, observamos, como se esperava, uma tendência crescente e um padrão de sazonalidade anual fortemente pronunciado.

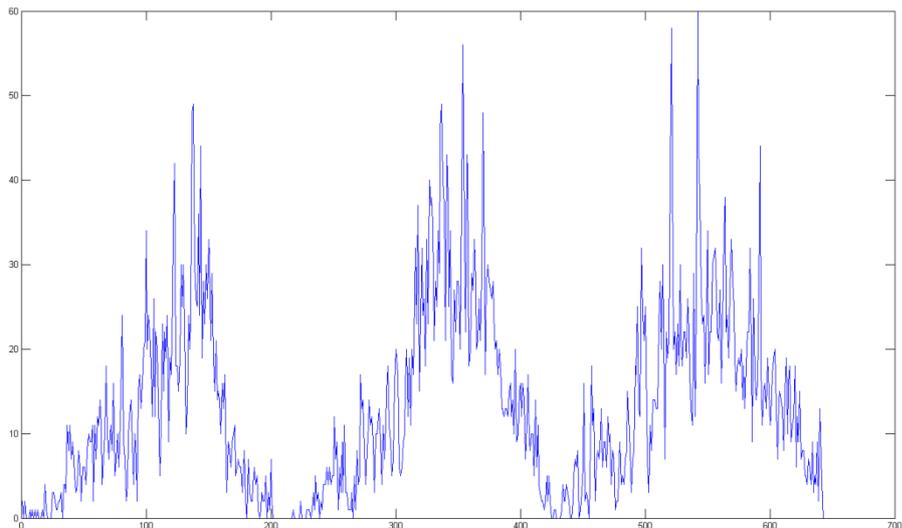


Figura 17- Série temporal de vendas diárias de calçado da categoria sandália de uma loja na estação Primavera/Verão de 2007, 2008 e 2009

O gráfico da respectiva função de auto-correlação amostral (Figura 18) mostra que:

- a série é não estacionária (os valores da auto-correlação mantêm-se elevados e positivos)
- a série é sazonal (os valores da auto-correlação para os lags 7, 14, 21 são sempre superiores às auto-correlações adjacentes) sendo a sazonalidade semanal(7 dias).

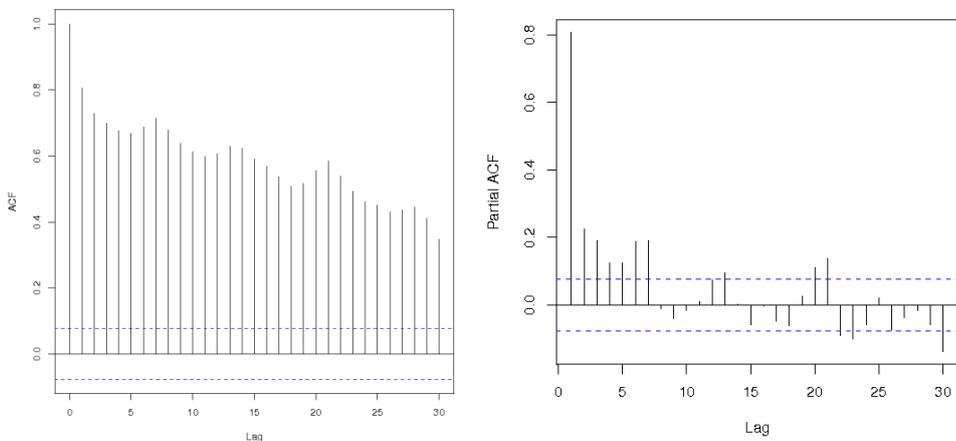


Figura 18 – Função de auto-correlação amostral e função de auto-correlação parcial amostral da série diária

O gráfico da função de auto-correlação parcial amostral confirma a sazonalidade semanal com picos até ao lag 7, e valores em torno de zero a partir desse lag. Pode assim concluir-se que no padrão de treino de entrada da RNA deverão ser incluídas as quantidades vendidas dos últimos 7 dias.

O gráfico da Figura 19 mostra o padrão de compra semanal relativo à mesma série. Pode observar-se que os valores de vendas mais elevados registam-se durante o fim-de-semana, e que nos dias úteis, os valores das quantidades vendidas vão crescendo ao longo da semana.

Este padrão evidencia que a informação sobre o dia da semana é relevante para o processo de aprendizagem da rede e portanto esta deverá ser também incluída no padrão de treino de entrada da RNA.

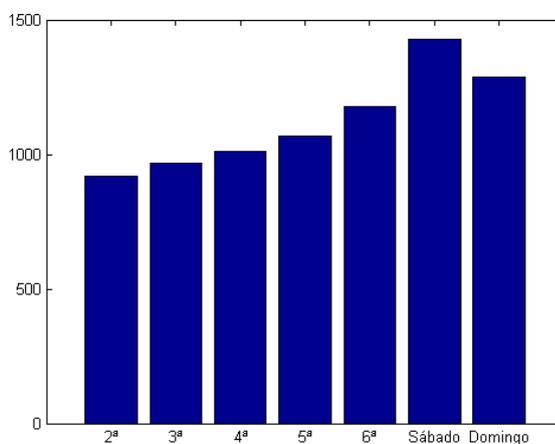


Figura 19 – Padrão de compra semanal.

Com o intuito de desenvolver um modelo de previsão causal foi inspeccionada a relação entre as quantidades vendidas diariamente e certas variáveis consideradas “explicativas”. O gráfico da Figura 20 mostra a relação entre as vendas diárias de calçado da categoria sandália na estação Primavera/Verão de 2007, 2008 e 2009 e a temperatura média diária observada no mesmo período.

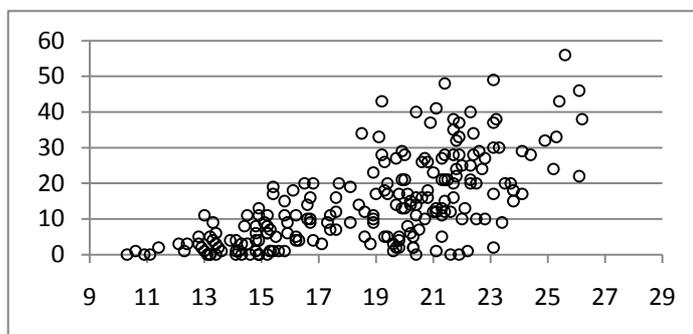


Figura 20 – Vendas diárias de calçado da categoria sandália de uma loja na estação Primavera/Verão de 2007, 2008 e 2009 vs temperatura média.

Tal como se esperava, as quantidades vendidas são afectadas pela temperatura, estando as temperaturas mais baixas associadas a quantidades vendidas baixas e temperaturas mais elevadas associadas a quantidades vendidas também elevadas. Conclui-se assim que o valor da temperatura média diária é relevante para o processo de aprendizagem da rede e portanto esta variável deverá ser incluída no padrão de treino de entrada da RNA.

No gráfico da Figura 21 mostra-se o padrão mensal do índice de volume de negócios no comércio a retalho de têxteis, vestuário, calçado e artigos de couro entre Janeiro de 2007 e

Dezembro de 2009. Este gráfico mostra um padrão de sazonalidade anual que é naturalmente relevante para a previsão em causa, pelo que se concluiu que esta variável também deveria ser incluída no padrão de treino de entrada da RNA.

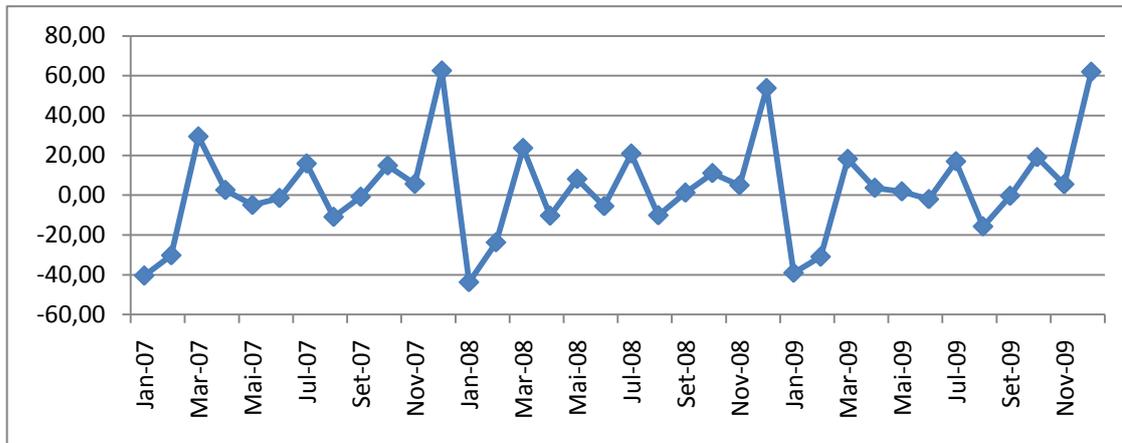


Figura 21 – Padrão mensal do índice de volume de negócios no comércio a retalho de têxteis, vestuário, calçado e artigos de couro entre Janeiro de 2007 e Dezembro de 2009 (Fonte: site do INE).

Uma inspeção detalhada e uma análise exploratória às séries de vendas diárias revelou que o período de saldos, os períodos de promoções e eventos especiais, os feriados, as semanas pré e pós Natal/Páscoa (referentes às férias escolares) e a disponibilidade de artigos/modelos em loja, afectam de forma significativa as quantidades vendidas pelo que se concluiu que esta informação também deveria ser incluída no padrão de treino de entrada da RNA.

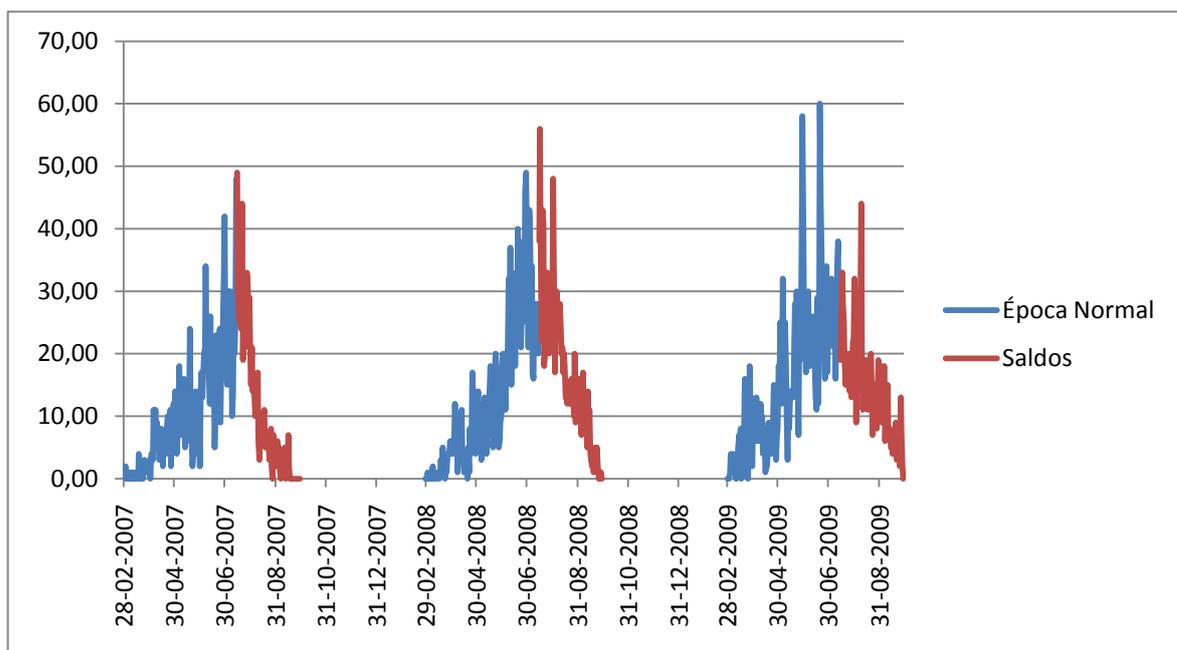


Figura 22 - Vendas diárias de calçado da categoria sandália de uma loja na estação Primavera/Verão de 2007, 2008 e 2009. Época normal e época de saldos

4.4.2.-Definição dos *inputs*

As variáveis que se mostraram determinantes para o modelo de previsão causal das vendas diárias de calçado, e que constituíram o padrão de treino de entrada da RNA (matriz de input) foram as seguintes:

- Mês (variável *dummy*)
- Dia da semana (variável *dummy*)
- Feriado (variável *dummy*)
- Promoções e eventos especiais (variável *dummy*)
- Período de saldos (variável *dummy*)
- Semana pré e pós Natal/Páscoa (variável *dummy*)
- Temperatura média (estação Primavera/Verão)
- Índice de volume de negócios no comércio a retalho de têxteis, vestuário, calçado e artigos de couro (INVCR)
- Vendas diárias dos últimos 7 dias corrigidas de tendência e sazonalidade (y_{t-7} , y_{t-6} , y_{t-5} , y_{t-4} , y_{t-3} , y_{t-2} e y_{t-1})

A matriz de *target* foi constituída pelas vendas diárias (y_t).

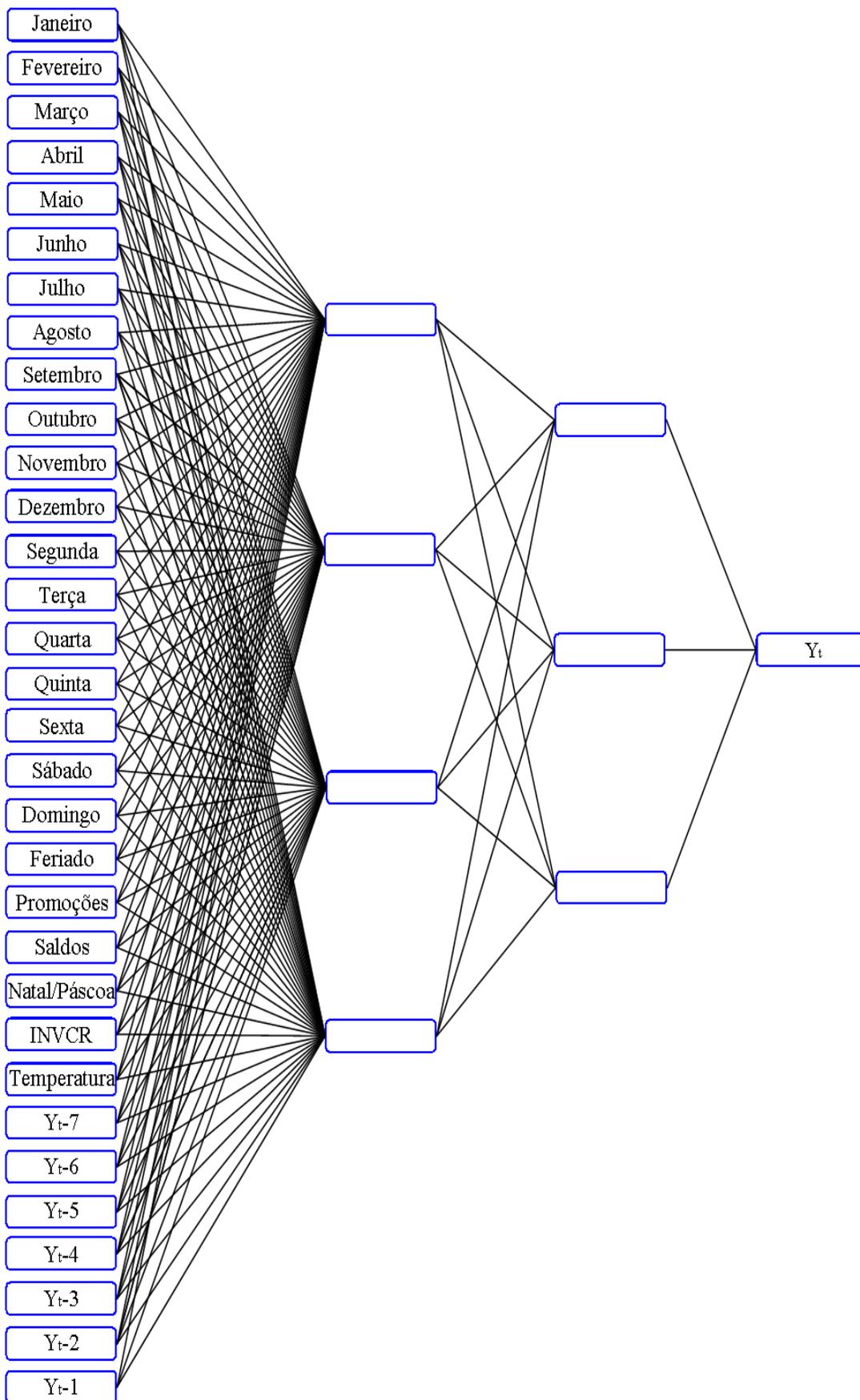


Figura 23 – estrutura da RNA utilizando os inputs propostos

4.4.3.- Resultados

Foram feitos testes com várias selecções de inputs, entre as quais foram testadas a performance da rede com e sem os inputs de temperatura e do INVCR. Estes testes foram feitos com o objectivo de tentar escolher a rede que apresentava menores erros.

Tabela 6 – resultados dos erros de previsão diárias utilizando variáveis explicativas.

Inputs	Rede	Arquitectura da Rede	EQM	REQM	EAM
Temperatura média	1	25-4-3-1	51.4577	7.1734	5.6176
INVCR	2	25-4-3-1	96.00	9.7983	6.6439
INVCR+temperatura média	3	26-4-3-1	64.2835	8.0177	6.3026

De entre as simulações feitas com as vários inputs foi verificado que a RNA apresenta piores resultados quando a variável INVCR é inserida nos *inputs*. Este resultado pode ser explicado por não haver grande correlação entre a série temporal de vendas e o INVCR. Por outro lado a inserção da temperatura média melhora a performance da rede e aliás torna-a na rede com melhores resultados.

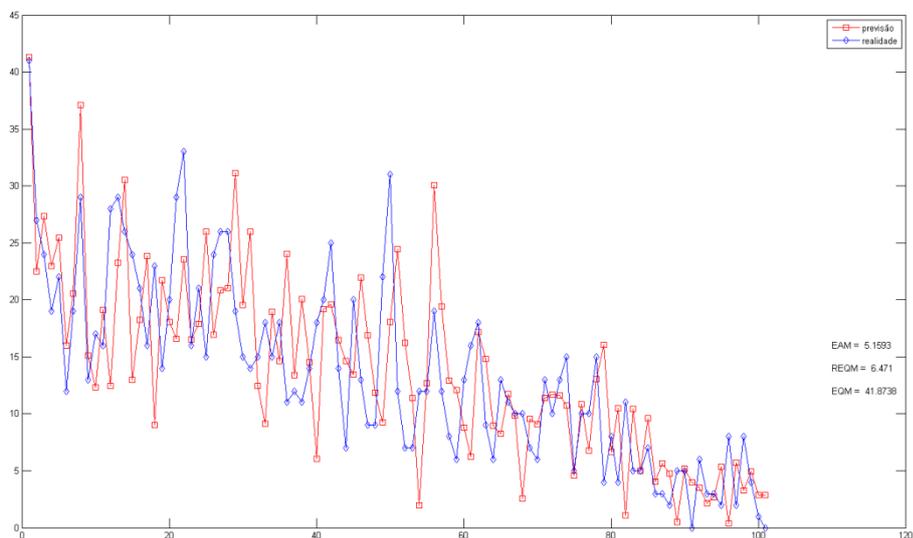


Figura 24 – Previsão vs Realidade da loja que mais vende

Aplicação de Redes Neurais na Previsão de Vendas para Retalho

Foram feitas 3 amostras de 50 simulações cada para verificar se os erros da rede 1 se mantinham, e não andavam a flutuar em relação á primeira simulação.

Tabela 7- resultados das 3 amostras de 50 simulações cada

Rede 1		Amostra 1		
	Mínimo	Máximo	Média	
REQM	6,4186	9,06392	7,37217	
EQM	41,1985	82,1546	54,9792	
EAM	4,69527	6,83584	5,6427	

Rede 1		Amostra 2		
	Mínimo	Máximo	Média	
REQM	6,4186	10,2875	7,53871	
EQM	41,1985	105,832	57,5987	
EAM	4,69527	7,32921	5,76493	

Rede 1		Amostra 3		
	Mínimo	Máximo	Média	
REQM	6,4186	10,2875	7,53871	
EQM	41,1985	105,832	57,5987	
EAM	4,69527	7,32921	5,76493	

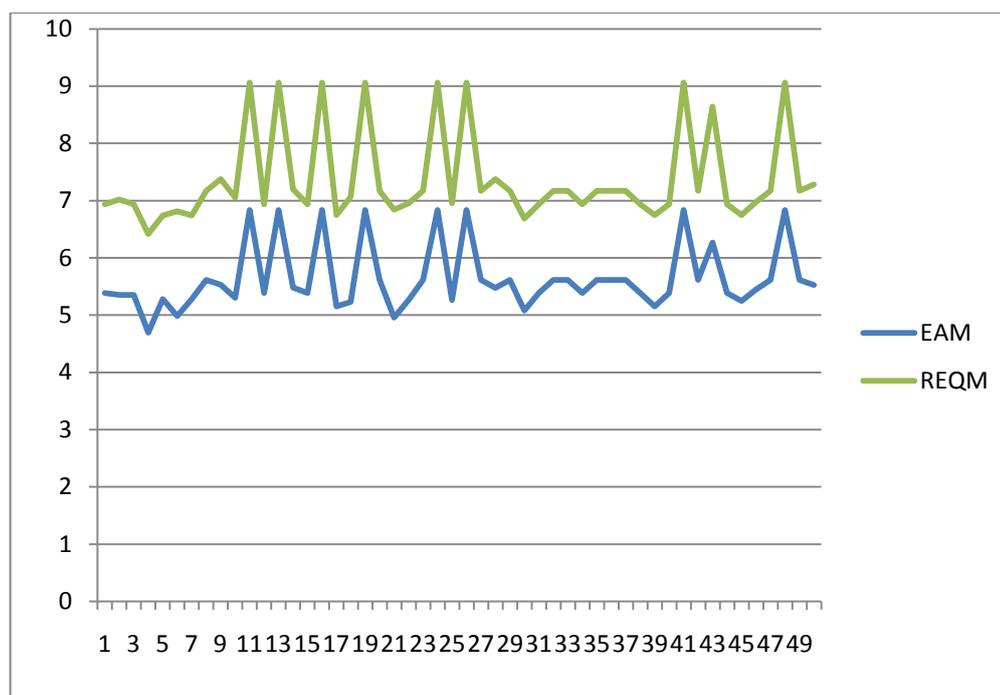


Figura 25 EAM e REQM da amostra 1

Com a análise da tabela 7 e figura 20 podemos verificar que nas 150 simulações a rede 1 tem um EAM médio de 5.6114 sapatos, um mínimo de 4.6953 sapatos e um valor máximo de 7.3292 sapatos. Também se obtém os valores para o REQM de média 7.3303 sapatos, um mínimo de 6.4186 sapatos e um máximo de 10.2875 sapatos. Podemos concluir a partir desta análise que a performance da rede 1 está situada sempre na mesma gama de valores, o que nos indica a capacidade de generalização que esta possui e também que com estes *inputs* as previsões efectuadas vão estar sempre à volta dos mesmos valores e não com grandes flutuações.

4.5.- Previsão de subcategorias

Após ter sido feita uma previsão diária bem sucedida para a categoria mais vendida da empresa, o passo seguinte foi verificar se, com a RNA que nos trouxe melhores resultados na previsão da categoria, também era possível fazer uma previsão um nível abaixo, no caso, previsão de vendas de subcategoria.

A subcategoria escolhida foi a mais representativa da categoria e correspondia a 30% das vendas da mesma.

Tabela 8 – Previsão de subcategoria utilizando a rede 1

Rede	Arquitectura da Rede	EQM	REQM	EAM
1	25-4-3-1	51.4577	7.1734	5.6176

4.6.-Capacidade de generalização da rede

Com o intuito de averiguar se a RNA treinada com os dados da loja que mais vende tem capacidade de fazer uma previsão de outras lojas que tenham um perfil de vendas similar, foi feita uma análise de correlação entre as séries temporais de vendas de todas as lojas. Na tabela mostramos a quantidade de correlações, consideradas aceitáveis, que essa loja tem.

Tabela 9 – Análise de correlação das séries temporais de todas as lojas

Correlação > 0.6	Correlação > 0.7	Correlação > 0.8	Correlação > 0.9
28 lojas	20 lojas	5 lojas	0 lojas

Foram escolhidas 3 lojas de entre o universo de mais de 50 que constituem a empresa, uma com correlação entre 0.6 e 0.7, outra com correlação entre 0.7 e 0.8 e finalmente a loja que apresentava maior relação com a loja que mais vendia.

Utilizando a rede 1, pois foi a rede que apresentou melhores resultados, treinada com dados da maior loja, tentou-se fazer a previsão das lojas escolhidas.

Analisando a tabela 10 percebemos que as previsões feitas nas redes treinadas com os dados das próprias lojas correlacionadas com a loja que mais vende, são muito similares com as previsões feitas com a rede treinada com os dados da rede da maior loja, no entanto a previsão feita para a loja com correlação 0.6 tem uma performance um pouco pior, mas a diferença de performance não justifica fazer uma RNA para as lojas que apenas possuam uma correlação de 0.6, pois deste modo com apenas a rede 1 conseguiremos prever 28 lojas. Estes resultados permitem concluir que a rede tem boa capacidade de generalização e de identificação de padrões. Foi concluído também que com a rede 1 podemos fazer a previsão de vendas de qualquer loja da empresa que apresente uma série temporal com pelo menos 0.6 de correlação com a série temporal de vendas da maior loja, este facto é uma mais valia, pois ,mais tarde aquando da implementação do método de previsão utilizando RNAs na empresa, será possível utilizar poucas RNAs para prever as vendas globais da empresa, o que acarretará menores encargos, pois menos RNAs implicam:

- Menor tempo de treino;
- Menor tempo a efectuar previsões;

Tabela 10 comparação da performance das previsões feitas com maior rede vs rede da própria loja

	EQM rede 1	REQM rede 1	EAM rede 1	EQM rede da própria loja	REQM rede da própria loja	EAM rede da própria loja
Correlação > 0.6	34.9696	5.9135	4.5085	18.071	4.251	3.3582
Correlação > 0.7	13.9257	3.7317	2.717	15.4742	3.9337	2.8509
Correlação > 0.8	29.48	5.4296	4.0206	30.0545	5.4822	4.0539
Correlação > 0.9	-	-	-	-	-	-

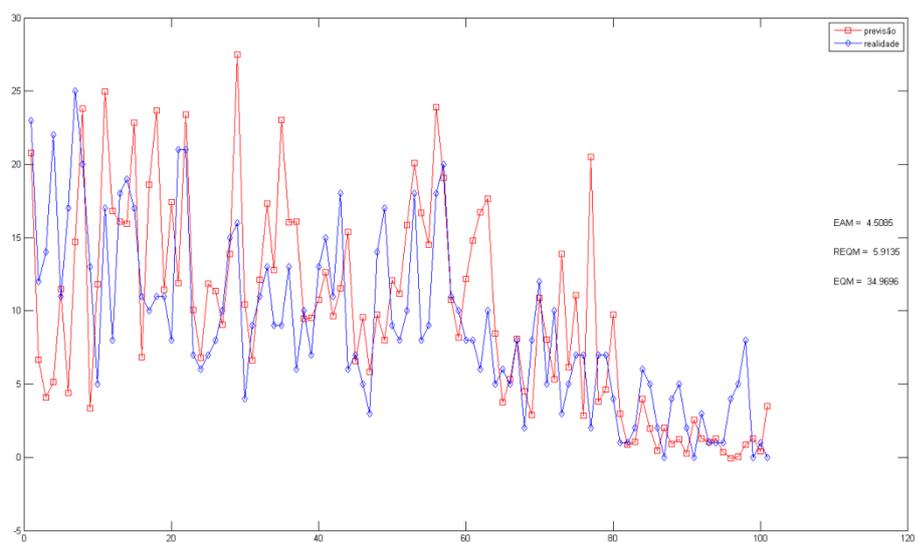


Figura 26 - Previsão vs Realidade de uma loja com correlação 0.6 > prevista utilizando a rede treinada da loja que mais vende

Aplicação de Redes Neurais na Previsão de Vendas para Retalho

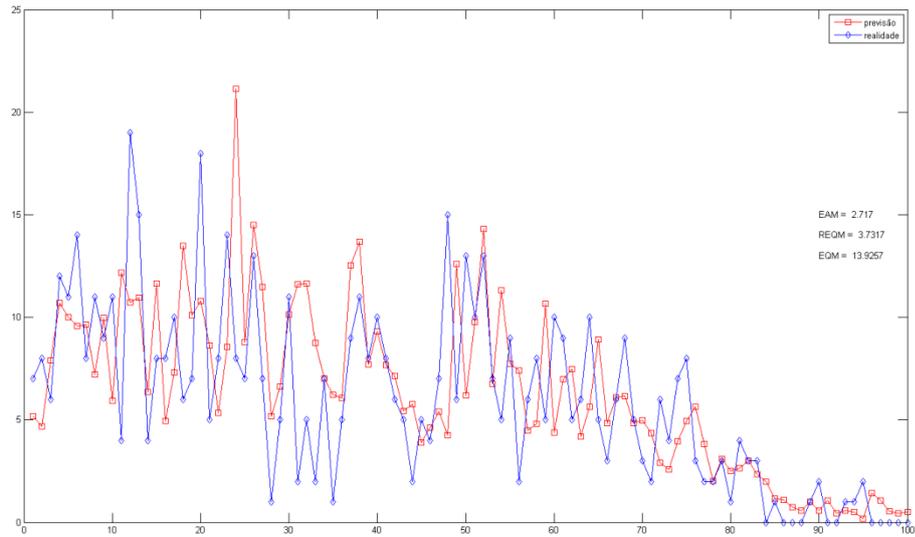


Figura 27- Previsão vs Realidade de uma loja com correlação > 0.7 prevista utilizando a rede treinada da loja que mais vende

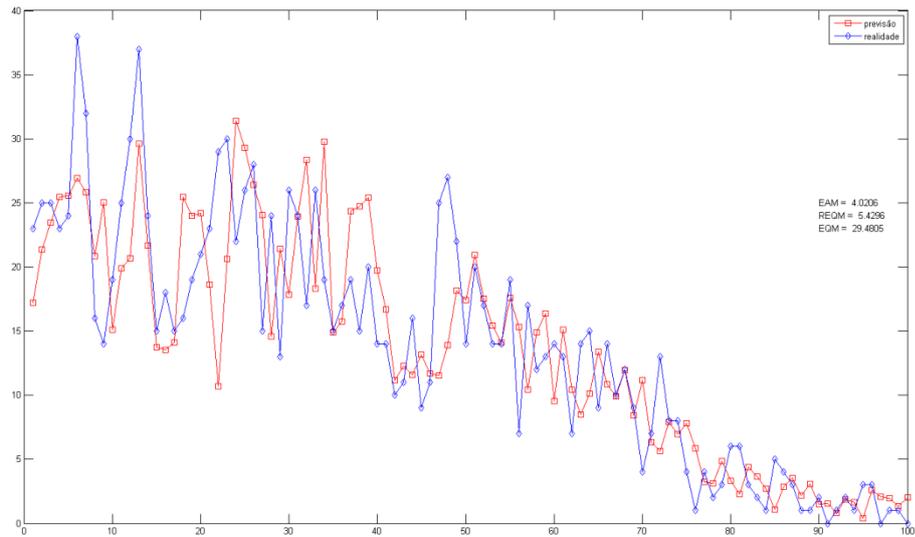


Figura 28- Previsão vs Realidade de uma loja com correlação > 0.8 prevista utilizando a rede treinada da loja que mais vende

5.- Conclusão

5.1 - Conclusões

Esta dissertação centrou-se na análise da aplicabilidade das RNAs na previsão de vendas para retalho. Não existe a pretensão de demonstrar que as RNAs são a melhor ferramenta para efectuar previsão de futuros valores de séries temporais, até porque, não foi feita comparação com outros modelos de previsão, mas pode ser dito com segurança que as RNAs são uma ferramenta poderosa para este tipo de problema.

Nesta dissertação foram exploradas as RNAs do tipo *feedforward* treinadas com algoritmo de retropropagação de erros para a previsão de séries temporais de vendas na indústria do calçado. De acordo com os resultados obtidos, a performance foi positiva, tendo-se concluído que as RNAs podem facilmente e com boa precisão fazer uma previsão, apesar dos dados terem ruído, serem algo aleatórios e apresentarem não-linearidades por vezes difíceis de capturar.

Este método de utilização da RNA pode ser de grande importância para a indústria de retalho, dado que assiste nas decisões. Isto é especialmente verdade uma vez que nas vendas ao retalho as tendências e a sazonalidade nas séries temporais são muitas vezes perturbadas quando factores como as condições meteorológicas e promoções são considerados. Mesmo nestas condições a RNA apresenta uma boa performance.

Embora o modelo possa ser melhorado recorrendo ao uso de mais variáveis explicativas que tragam benefícios em entender a série temporal e também mais dados, de modo a que a RNA minimize o erro, servindo-se da capacidade de aprender mais padrões.

Uma das dificuldades deste problema foi o facto de se estar a fazer previsão de um produto do segmento moda. Este é um segmento que acarreta particulares dificuldades de previsão dado a constante mudança das tendências de moda. No entanto o uso das RNAs neste problema em específico pode ser amplificado com a ajuda de um especialista capaz de avaliar as previsões de vendas em função das novas tendências da moda que consiga dar uma previsão qualitativa que aliada à previsão quantitativa da RNA possa minimizar o erro de previsão da mesma.

Um dos grandes contributos deste estudo é o desenho do processo a seguir para, com o auxílio das RNAs, efectuar uma previsão de vendas, com base nos dados da empresa, ou seja, fica definida uma metodologia capaz de aliar a teoria à prática e permitir a utilização efectiva destas ferramentas. Como suporte à aplicação desta metodologia, será desenvolvido um software com um ambiente *user-friendly* que simplificará todo o processo.

Através do trabalho desenvolvido identificou-se que a origem das rupturas está associada a problemas de distribuição. A RNA óptima pode ser utilizada para prever as vendas diárias de calçado e essa informação pode ser processada para ajudar ao planeamento das operações da empresa. Estimou-se que, no cenário mais conservador, existe um potencial de crescimento das vendas na ordem dos 10% por estação. A previsão de vendas por loja/artigo, a curto prazo, deve funcionar como input ao algoritmo de distribuição entre armazém e loja com o intuito de colocar o produto certo, no local certo no tempo certo.

A grande dificuldade que se apresentou na resolução deste problema de previsão utilizando RNAs passou pela configuração das RNAs. Para definir a rede a ser utilizada, foi necessário escolher o algoritmo de treino, número de *layers*, neurónios de cada *layer*, função de transferência, tipo de normalização dos dados e divisão dos dados para treino e validação. O facto de ainda não existir uma metodologia capaz de apresentar a melhor configuração da rede para cada série temporal, levou a que este processo se tornasse iterativo o que o tornou extremamente moroso.

A empresa deve avaliar os custos e benefícios das RNAs antes de escolher a devida ferramenta de previsão. Os custos deverão ser avaliados tendo em conta:

- A quantidade e qualidade dos dados;
- O nível técnico dos peritos;
- O software específico;
- O elevado tempo de computação necessário.

5.2.- Futuros Desenvolvimentos

Para continuação do trabalho realizado nesta dissertação, sugerem-se os seguintes desenvolvimentos:

- Explorar metodologias de análise de sensibilidade para averiguar a utilidade das variáveis explicativas utilizadas;
- Comparação dos resultados obtidos com outros métodos de previsão ;
- Implementação de um modelo de previsão desenvolvido numa aplicação de software para ser aplicada por parte da empresa;
- Investigar novas variáveis explicativas, tais como custo da colecção, número de modelos disponíveis, concorrentes directos no mercado, qualidade da colecção e qualidade da colecção dos concorrentes;
- Investigação de modelos RNA com poda automática de neurónios que apresentem menor importância durante o treino da rede.

Referências Bibliográficas

- Alon I., Qi M. and Sadowski R. J. (2001). Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods. *Journal of Retailing and Consumer Services*, Volume 8, pp. 147-156.
- Chua C.-W. and Zhang G. (2003). A comparative study of linear and nonlinear models for aggregate retail sales forecasting. *International Journal of Production Economics*, Volume 86, Issue 3, pp. 217-231.
- Das P. and Chaudhury S. (2007). Prediction of retail sales of footwear using feedforward and recurrent neural networks. *Neural Computing & Applications*, Volume 16, Numbers 4-5, pp. 491-502.
- Nelson M., Hill T., Remus T. and O'Connor, M. (1999). Time series forecasting using neural networks: Should the data be deseasonalized first? *Journal of Forecasting*, Volume 18, pp. 359-367.
- Reed R. D. and Marks R. J. (1999). *Neural Smoothing: Supervised Learning in Feedforward Artificial Neural Networks*. MIT Press.
- Smith K. and Gupta J. (2002). *Neural Networks in Business: Techniques and Applications*. Idea Group Publishing.
- Tang Z., de Almeida C. and Fishwick, P. A. (1991). Time series forecasting using neural networks vs. Box-Jenkins methodology. *Simulation*, Volume 57, Issue 5, pp. 303-310.
- Zhang G., Patuwo B. E. and Hu M. Y. (1998). Forecasting with artificial neural networks: The state of the art, *International Journal of Forecasting*. Volume 14, Issue 1, pp. 35-62.
- Zhang G. (2004). *Neural Networks in Business Forecasting*. Idea Group Publishing.
- Zhang G. and Qi M. (2005). Neural network forecasting for seasonal and trend time series. *European Journal of Operational Research*, Volume 160, Issue 2, pp. 501-514.
- Lippmann, R.P., (1987). An introduction to computing with neural nets, *IEEE ASSP Magazine*, April, 4-22.
- Chakraborty, K., Mehrotra, K., Mohan, C.K., Ranka, S., (1992) Forecasting the behavior of multivariate time series using neural networks. *Neural Networks* 5, 961-970.
- Cybenko, G., 1988. Continuous Valued Neural Networks with Two Hidden Layers are Sufficient. Technical Report, Tuft University.
- Geman, S., Bienenstock, E. and Doursat, (1992) R. Neural networks and the bias/variance dilemma. *Neural computation*, no. 4, pp. 1-58.
- Park, J., Sandberg, I.W., (1991). Universal approximation using radial basis function networks. *Neural Computation* .
- Au, Kin-Fan e Choi, Tsan-Ming. (2008). Fashion retail forecasting by evolutionary neural networks. *International Journal of Production Economics*. Volume 114.
- Armstrong, Scott. (2001). *Principles of Forecasting: A Handbook for Researchers and Practitioners*. Kluwer Academic Publishers.

Aplicação de Redes Neurais na Previsão de Vendas para Retalho

Caudill, M. (1990). "Neural networks primer, Part III." *AI Expert*, 3(6) .