

**UNIVERSIDADE FEDERAL DE LAVRAS  
DEPARTAMENTO DE CIÊNCIAS EXATAS (DEX)**

Orientador: Prof. Fabiano Lemes Ribeiro

**Estudo de Algoritmos de Aprendizado em  
Redes Neurais Artificiais**

Nome: João Vitor Squillace Teixeira

Relatório Parcial apresentado à  
Universidade Federal de Lavras, como parte das  
exigências do PIBIC/CNPq ou FAPEMIG ou PIVIC,  
referente ao período de março/2012 a março/2013.

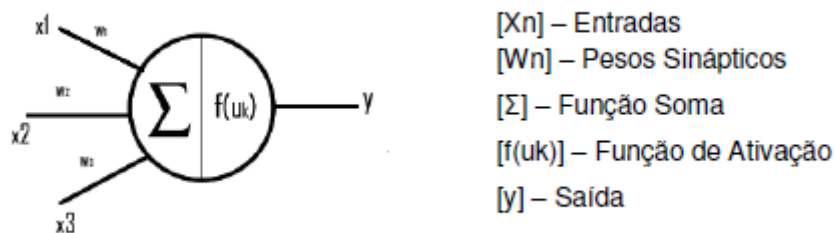
**LAVRAS/MG – NOVEMBRO/2012**

São descritos neste relatório os estudos realizados utilizando redes neurais artificiais para a resolução de problemas de classificação e previsão de dados. Para a classificação a rede neural será utilizado uma arquitetura Single-Layer implementada a partir do algoritmo de Hebb, e para a previsão será utilizado a arquitetura Multi-Layer com algoritmo Backpropagation. Nas pesquisas também foram abordados estudos sobre evolução do cérebro e dos sistemas neurais que serão citados ao fim da introdução.

## 1. Introdução

As redes neurais são formadas por diversos elementos que tentam imitar o funcionamento do dos neurônios. De forma análoga ao cérebro humano, as redes adquirem conhecimento através de um ambiente externo e tentam encontrar a solução para determinado problema através de um processo de aprendizado.

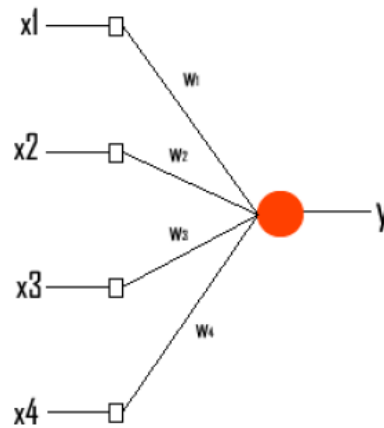
As redes neurais são compostas por neurônios artificiais, estes funcionam como processadores dos dados recebidos. Eles são formados basicamente por três elementos de base: pesos sinápticos, função de soma e função de ativação conforme mostra a figura 1.1.



**Fig 1.1 Neurônio Artificial**

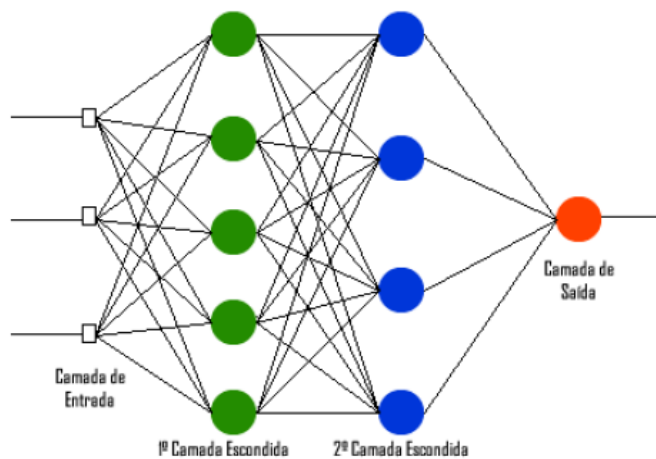
As redes neurais com arquitetura *Single-layer Feedforward* são redes mais simples, que possuem apenas uma camada de entrada ligada diretamente a uma camada de saída. A rede é estritamente progressista, ou seja, nesse tipo de arquitetura os dados vão em direção apenas da camada de entrada para a camada de saída.

Conforme mostra a figura 1.2, os dados de entrada são compostos pela representação  $X_n$ , os pesos sinápticos são representados por  $W_n$  e a função de ativação é representada pela bola laranja.



**Fig 1.2 Single-layer**

As redes com arquitetura *Multi-layers* possuem camadas intermediárias chamadas de camadas ocultas ou escondidas. As funções das camadas escondidas são de intervir entre as camadas de entrada e saída de modo a produzir a diminuição do erro, nesse tipo de arquitetura podemos ter algoritmos que propagam os dados não só no sentido entrada-saída, mas também no sentido saída-entrada.



**Fig 1.3 Multi-layer**

As pesquisas até o momento, além de abordar o desenvolvimento das técnicas para resolução de problemas utilizando redes neurais, também ocorre a partir da absorção da literatura na grande área da neurociência com o intuito de melhor relacionar as redes neurais artificiais com as redes neurais biológicas. Com isso o estudo se inicia com a evolução do sistema nervoso desde os primórdios da vida até os dias atuais, e o peso social das funções cerebrais decorrente dessa evolução.

## 2. Objetivos

Na utilização da arquitetura *single-layer*, o problema consiste na classificação de um óleo em duas classes de pureza a partir de três de suas propriedades físico-químicas. Já na utilização da arquitetura *multi-layer*, é implementado uma rede neural capaz de fazer a previsão da bolsa de valores para o ano de 2011.

## 3. Metodologia

Para as implementações das redes foram utilizadas o software Scilab. Este é um software científico computacional, que apresenta um toolbox chamado ANN com funções voltadas especialmente para redes neurais.

### 3.1. Classificação das classes de pureza do óleo

Os dados de entrada do problema de classificação eram constituídos de uma amostra de 30 dados para treinamento e 10 dados para teste. Foi utilizado para o treinamento o algoritmo do *perceptron* empregando a regra de Hebb, onde os ajustes dos pesos das conexões são dados pelo produto da atividade pré-sináptica e pós-sináptica (eq.3.1.1) passando por uma função de ativação  $g$  (eq.3.1.2).

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(t). \text{ (eq.3.1.1)}$$

$$\Delta w_{ij}(t) = g(y_i(t), x_j(t)) \text{ (eq.3.1.2)}$$

Inicialmente, todos os dados são armazenados em um vetor e posteriormente separados através de um laço em dados de treinamento e

dados de teste. Os pesos são iniciados aleatoriamente para que a rede possa ser iniciada. A partir das inicializações a função de ativação é chamada de modo a ativar ou não o peso atual, após isso é calculado o erro pela expressão (eq.3.1.3), onde  $e(i)$  representa o erro atual,  $yd(i)$  a saída calculada e  $y(i)$  é a saída real.

$$e(i) = yd(i) - y(i) \text{ (eq.3.1.3)}$$

Depois de calculado o erro, é feito por fim os ajustes dos pesos sinápticos a partir da expressão (eq.3.1.4), onde  $w$  são os vetores dos pesos,  $\eta$  é a taxa de aprendizagem da rede,  $e(i)$  o erro atual calculado e  $x(i)$  é o dado de entrada atual da rede.

$$w(t+1) = w(t) + (\eta * e(i) * x(i)) \text{ (eq.3.1.4)}$$

Os ajustes dos pesos são internos a um laço que possui duas condições, a primeira representando as épocas\* (pré-determinadas) da rede ou a segunda, onde erro alcança um valor mínimo, definido como 0,01.

\*As épocas são as quantidades de vezes que os dados de treinamento são apresentados a rede.

## 3.2. Previsão do índice da bolsa de valores

Para o problema de previsão foram utilizados os dados disponíveis no site da BMF Bovespa (<http://www.bmfbovespa.com.br>).

Na previsão foram utilizadas as cotações diárias do fechamento do Ibovespa no período de janeiro de 1998 até dezembro de 2011 totalizando 3463 dados. A divisão dos dados foi de aproximadamente 92,83% (3215 dados) de treinamento e 7,16% (248 dados) para teste, resultando apenas na previsão do ano de 2011.

O algoritmo de treinamento utilizado foi o *backpropagation*, já presente no toolbox ANN de redes neurais do Scilab. O algoritmo *backpropagation* é basicamente realizado em duas etapas. Na primeira etapa, os dados são propagados pela rede no sentido da entrada para a saída, chamada de fase *forward*, sem que ocorram quaisquer alterações nos pesos. Já na segunda a resposta da rede é comparada com a resposta desejada e os pesos são reajustados de maneira a minimizar o erro, esta chamada de fase *backward*.

Nesta segunda etapa o ajuste de um peso  $w$  que define seu valor para a próxima iteração é dado pela expressão (eq.3.2.1).

$$w(t+1) = w(t) + (\eta * \delta(i) * x(i)) \text{ (eq.3.2.1)}$$

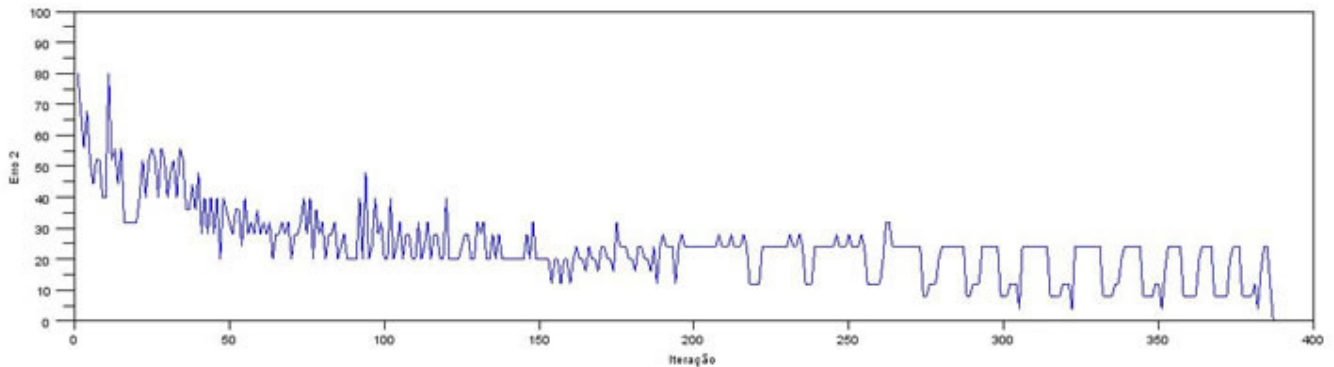
Onde na (eq.3.2.1)  $\eta$  é a taxa de aprendizagem e  $\delta$  é o gradiente local do erro para o neurônio. O gradiente local do erro  $\delta$  é determinado através do método do gradiente descendente, que não será abordado no relatório em questão.

## 4. Resultados e Discussão

Para o problema de classificações foi obtido os seguintes resultados  $y$ , dado os dados de entrada para teste  $x_1$ ,  $x_2$  e  $x_3$ .

Amostra	$x_1$	$x_2$	$x_3$	$y$
1	-0.3565	0.0620	5.9891	-1
2	-0.7842	1.1267	5.5912	1
3	0.3012	0.5611	5.8234	-1
4	0.7757	1.0648	8.0677	1
5	0.1570	0.8028	6.3040	1
6	-0.7014	1.0316	3.6005	1
7	0.3748	0.1536	6.1537	-1
8	-0.6920	0.9404	4.4058	1
9	-1.3970	0.7141	4.9263	-1
10	-1.8842	-0.2805	1.2548	-1

O erro quadrático do treinamento segue na figura 4.1.



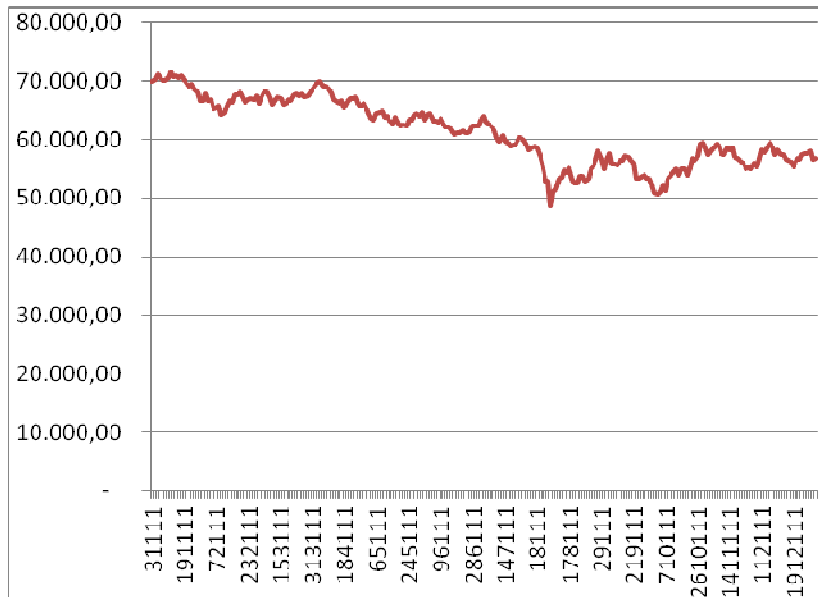
**Fig 4.1** Treinamento da rede a partir dos dados

O erro final do treinamento acabou convergindo à zero, isso se dá, pois os dados de treinamento eram linearmente separáveis, ou seja, os dados poderiam ser separados completamente de forma que resultassem em exatamente duas classes.

Para o problema de previsão, foi dada a rede os seguintes dados de entrada (fig. 4.2) e desejava-se prever o ano de 2011, onde o comportamento real do índice pode ser visto na figura 4.3.

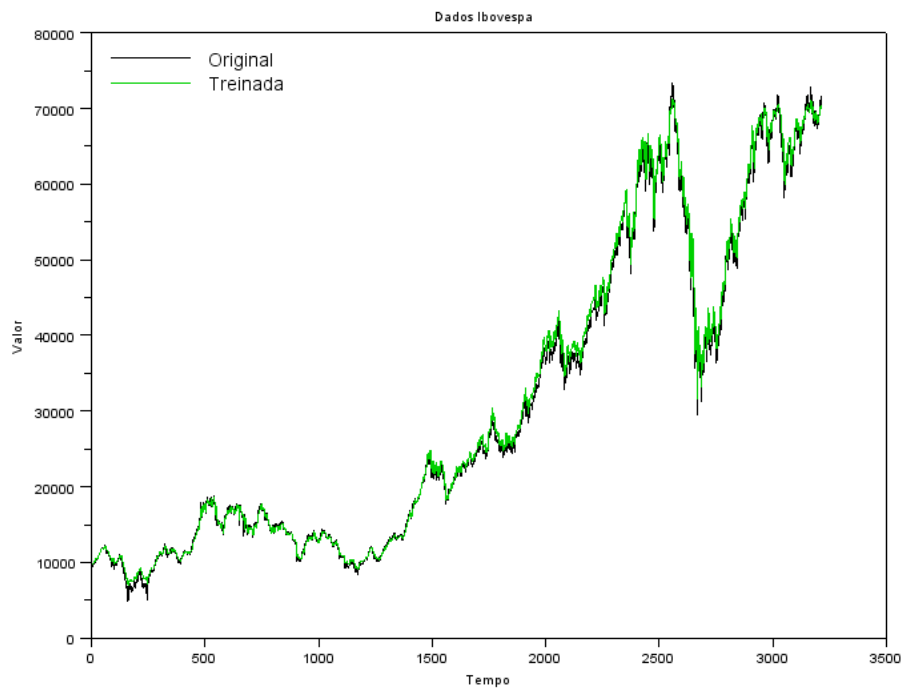


**Fig 4.2** Treinamento da rede a partir dos dados



**Fig 4.3** Treinamento da rede a partir dos dados

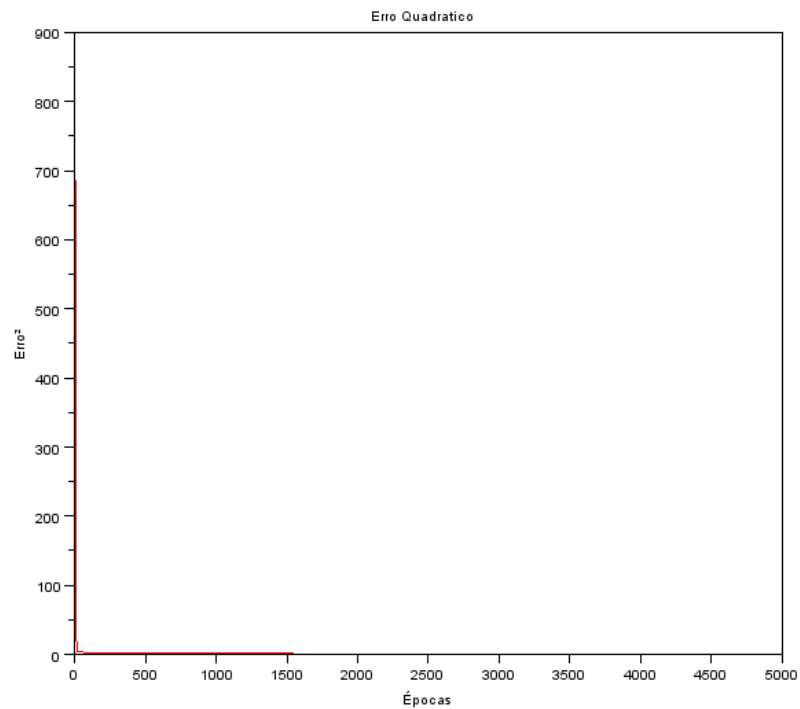
A figura 4.4 refere-se ao treinamento da rede que teve o menor erro de previsão.



**Fig 4.4** Treinamento da rede a partir dos dados

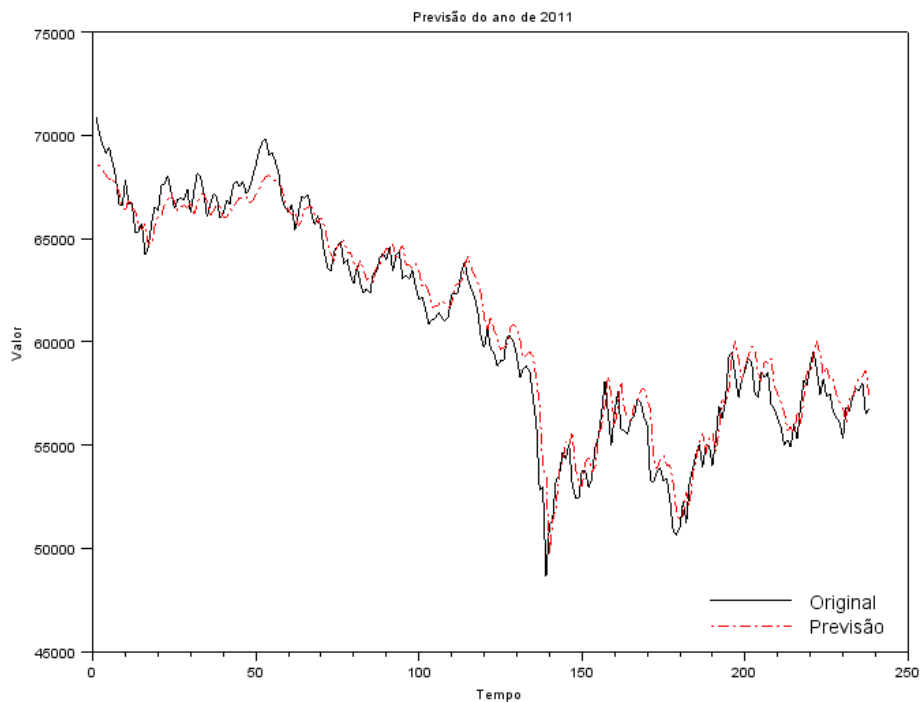


Na figura 4.5 pode-se observar o erro quadrático do treinamento em relação às épocas.



**Fig. 4.5 Erro quadrático do treinamento**

Para os testes foram utilizado os pesos sinápticos obtidos pelo treinamento, com variações de no máximo 2%. Na figura 4.6 segue a previsão da rede em relação às cotações reais do ano de 2011.



**Fig 4.6 Dados do Ibovespa 2011 e previsão da rede**

Uma rede neural funciona a partir de parâmetros dados, como por exemplo, taxa de aprendizado, momento ou até mesmo quantidade de neurônios nas camadas. Podemos ver isso claramente nos resultados da previsão, pois variando os parâmetros da rede se obtinha melhores ou piores previsões.

## 5. Conclusão

Os resultados obtidos revelam que as redes neurais artificiais, tanto no problema de classificação quanto no problema de previsão, apresentam resultados de forma eficiente, uma vez que os erros foram relativamente baixos no problema de previsão, e chegando à zero no de classificação.

A maior dificuldade na construção de uma rede neural com arquitetura *multi-layer* é a definição dos seus parâmetros, citado anteriormente, pois para cada problema específico tem-se resultados diferentes. A partir do momento que os parâmetros são ajustados de modo que o erro seja o mais baixo possível, a rede está configurada e poderá ser utilizada para resolver problemas da mesma origem do treinamento.

O atual andamento das pesquisas mostra de modo qualitativo o uso das redes neurais para uma possível simulação de alguns comportamentos decorrentes das organizações das redes biológicas, e com isso, nos indicar quais fatores são chaves no decorrer da evolução do sistema neural.

## 6. Referências

ANDRADE, André Luiz Monézi, JUNIOR, Alfredo Löhr. **The Implications of the Neural Plasticity In The Process of the Memory and Learning.** (2005)

DALGALARRONDO, Paulo. **Evolução do cérebro. Sistema nervoso, psicologia e psicopatologia sob a perspectiva evolucionista** – Porto alegre: ArtMed (2011)

BRAGA, Antônio de Pádua; CARVALHO, André Ponce de Leon F. de; LUDEMIR, Teresa Bernarda. **Redes Neurais Artificiais: Teoria e Aplicações**, 2º ed. (2007).

CERQUEIRA ,Eduardo O.; ANDRADE João C.; POPPI Ronei J.; MELLO, Cesar. **Redes Neurais E Suas Aplicações Em Calibração Multivariada**, UNICAMP 2001

CASTRO, Cristiano Leite de. **Sistema Inteligente para Previsão on-line do Índice Ibovespa baseado em Redes Neurais Artificiais**, UFMG 2003

FARIA, Elisângela Lopes de; ALBUQUERQUE, Marcelo Portes; ALFONSO, Jorge Luis González. **Previsão do Mercado de Ações Brasileiro utilizando Redes Neurais Artificiais**, CBPF-NT- 002/2008

BISHOP, Christopher M. **Neural Networks for Pattern Recognition**, 1<sup>o</sup> ed. 2006