

Redes neurais auto-organizáveis (SONN)

Manoel Fernando Tenório

Wei-Tsih Lee

Resumo para a matéria Redes neurais I.

Adriano Martins Moutinho

Estrutura

- ▶ Tem base em um processo iterativo de escolha e aproximação com base em funções de segundo grau.
- ▶ Cada neurônio possui apenas duas entradas e uma saída, com função de transferência de segunda ordem.
- ▶ A rede se modifica, tendo estrutura variável. Podendo se organizar de formas diferentes a cada treinamento.

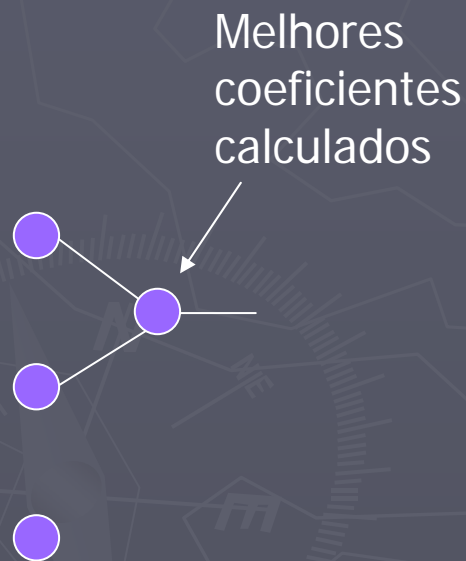
Funcionamento do algoritmo

- ▶ O algoritmo escolhe aleatoriamente, duas variáveis de entrada e combina as mesmas usando uma equação de segundo grau:

$$\hat{y} = a_0 + a_1x_1 + a_2x_2 + a_3x_1x_2 + a_4x_1^2 + a_5x_2^2$$

- ▶ Os coeficientes da equação são então calculados minimizando o erro em relação à saída, isto é feito segundo a seguinte equação.

Funcionamento do algoritmo (cont)



- ▶ Assim, a saída do novo neurônio é comparada com a saída desejada e uma função de conformidade é calculada. Caso a similaridade do novo neurônio seja melhor do que dos seus pais, o mesmo é aceito. Caso contrário, o algoritmo continua selecionando outras entradas.
- ▶ A saída do neurônio adicionado permanece e pode ser selecionada novamente pelo algoritmo

Funcionamento do algoritmo (cont)

- ▶ O calculo do melhor conjunto de coeficientes é feita com base na equação abaixo:

$$[Ent] \times [Coef] = [Saida]$$

- ▶ Onde a matriz de entrada é construída da seguinte forma, sendo as entradas z e w e n amostras:

$$[Ent] = \begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ w_1 & w_2 & w_3 & w_4 & \dots & w_N \\ z_1 & z_2 & z_3 & z_4 & \dots & z_N \\ w_1 \cdot z_1 & w_2 \cdot z_2 & w_3 \cdot z_3 & w_4 \cdot z_4 & \dots & w_N \cdot z_N \\ w_1^2 & w_2^2 & w_3^2 & w_4^2 & \dots & w_N^2 \\ z_1^2 & z_2^2 & z_3^2 & z_4^2 & \dots & z_N^2 \end{bmatrix}$$

Funcionamento do algoritmo (cont)

- ▶ Assim, isolando a matriz [coef]:

$$[Ent]^T \cdot [Ent] \cdot [Coef] = [Ent]^T \cdot [Saída]$$

$$([Ent]^T \cdot [Ent])^{-1} \cdot ([Ent]^T \cdot [Ent]) \cdot [Coef] = ([Ent]^T \cdot [Ent])^{-1} \cdot [Ent]^T \cdot [Saída]$$

$$[Coef] = ([Ent]^T \cdot [Ent])^{-1} \cdot [Ent]^T \cdot [Saída]$$

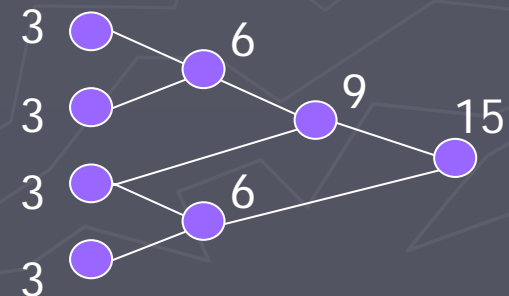
- ▶ Este processo escolhe os coeficientes que minimizam a diferença entre a saída e um polinômio de segundo grau gerado a partir de duas variáveis do sistema.

Funcionamento do algoritmo (cont)

- ▶ No entanto, é preciso também minimizar a estrutura utilizada na rede.
- ▶ Deseja-se não só minimizar o MSE mas também manter a estrutura o mais simples possível. Para isto, utiliza-se uma função de performance chamada SEC.
- ▶ Para tornar isto possível, as ligações buscam minimizar a função SEC abaixo, onde N é o número de padrões e k é gerado recursivamente somando os valores de k dos pais, sendo os nós primários inicializados com $k_0=3$:

$$SEC = \frac{1}{2} \cdot N \cdot \log(MSE) + \frac{1}{2} \cdot k \log N$$

Valores recursivos de K



Funcionamento do algoritmo (cont)

- ▶ Porém, é possível que uma determinada ligação entre duas variáveis não possua um SEC menor que o atual, mas isso poderia permitir uma melhor ligação no futuro.
- ▶ Para possibilitar melhores ligações futuras, utiliza-se um método baseado no processo de cura de metais.
- ▶ No processo de cura de metais, um determinado metal é esquentado até uma temperatura T e esfriado lentamente. Este processo permite que as ligações fortes entre as moléculas sejam formadas, embora inicialmente ocorram ligações fracas.

Funcionamento do algoritmo (cont)

- ▶ Seguindo o mesmo método, uma variável temperatura é inicializada com um valor alto, decrescendo com o passar do tempo. Dessa forma, enquanto a temperatura estiver alta, haverá maior probabilidade de aceitar um nó filho mesmo que não possua menor SEC que seus pais.
- ▶ Para isso, gera-se um número aleatório, que se for menor que $e^{(\text{delta}/\text{Temp})}$, o nó filho é aceito. Delta é a diferença entre o valor da menor energia e a energia atual, e Temp é a temperatura.

Funcionamento do algoritmo (cont)

- ▶ Como o número aleatório gerado está entre zero e um, caso a energia atual esteja próxima à energia mínima de todos os nós da rede, o nó filho tem muita chance de ser aceito. A probabilidade de aceitação também diminui conforme a temperatura decresce.
- ▶ Isso poderá possibilitar a criações de ligações mais fortes no futuro.
- ▶ Caso esta ligação não permita ligações mais fortes, a mesma será eliminada futuramente.

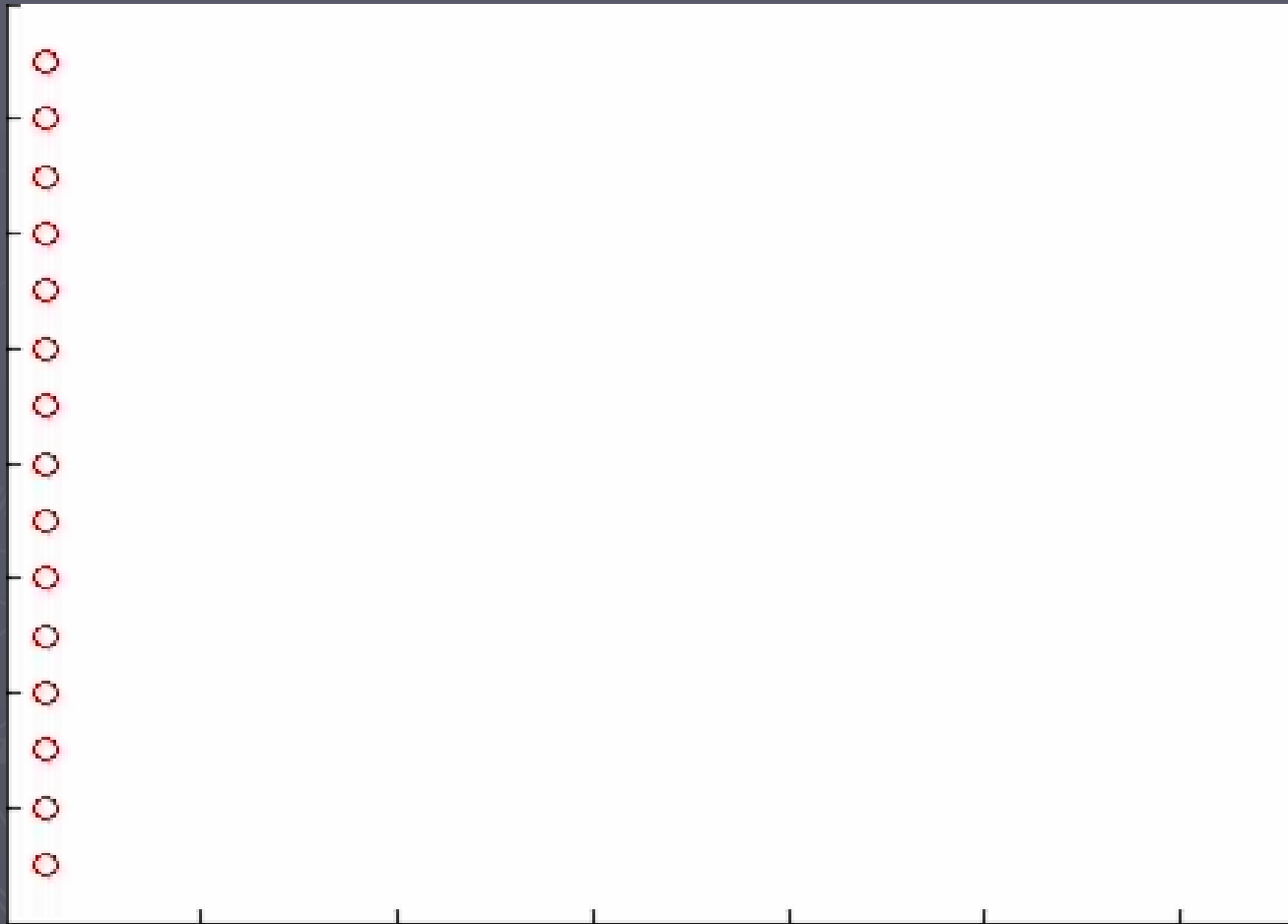
Resumo do algoritmo

- 1) Um neurônio para cada entrada é iniciado.
- 2) O algoritmo escolhe dois pares de entradas aleatoriamente.
- 3) Calcula-se os coeficientes para a aproximação polinomial de segunda ordem.
 - 1) Caso SEC desta aproximação seja melhor que os dois pais, este nó é adicionado.
 - 2) Caso não seja melhor, calcula-se um número aleatório, que se for maior que $e^{(\text{delta}/\text{Temp})}$ o nó é aceito mesmo assim.

Resumo do algoritmo (cont.)

- 4) Caso nenhum nó seja aceito, se $N/(N-1)$ nós já sejam tiverem sido escolhidos, baixa-se a temperatura de um valor estabelecido.
- 5) Caso nenhum nó seja aceito, volta-se para 2 sem baixar a temperatura.
- 6) Caso algum nó seja aceito por SEC, volta-se para 2 baixando-se a temperatura de um valor estabelecido.

Simulação da organização



IND: [2 3 4 7 11 18 19 31 33 34 35]

Resultados e testes

- ▶ Para testar a rede neural SONN, foi utilizado um problema de aproximação de funções cujo objetivo é estimar o fechamento da ação da Petrobrás no Ibovespa para três dias futuros, tendo-se o fechamento, máximo, mínimo, abertura e volume negociados nos dias anteriores.
- ▶ O modelo proposto SONN foi comparados com os tradicionais modelos de rede neural como MLP e a recorrente ELMAN.

Resultados e testes

MSE (MLP)		
Três dias	Dois dias	Um dia
0.048226	0.041824	0.019905
0.040887	0.033223	0.020943
0.054066	0.037466	0.03836
0.051453	0.04156	0.015988
0.042699	0.035581	0.024981
0.068195	0.044101	0.030539
0.055124	0.033703	0.028558
0.051975	0.030898	0.017736
0.061804	0.049583	0.029405
0.055558	0.040392	0.030278
Media (15 redes)	Media (15 redes)	Media (15 redes)
0.06572	0.04512	0.03259

Resultados e testes

MSE (ELMAN)		
Três dias	Dois dias	Um dia
0.047621	0.029378	0.021429
0.044724	0.027754	0.021564
0.03596	0.025407	0.017005
0.045063	0.034718	0.023135
0.040045	0.034424	0.025156
0.053318	0.039425	0.023397
0.044726	0.030707	0.023217
0.04389	0.032547	0.017709
0.046389	0.030483	0.021519
0.041549	0.029807	0.020426
Media (15 redes)	Media (15 redes)	Media (15 redes)
0.048346	0.03489	0.026571