

Fundamentos de Redes Neurais Artificiais

Victor São Paulo Ruela
Programa de Pós-Graduação em Engenharia Elétrica
Universidade Federal de Minas Gerais
Belo Horizonte, Brasil
Email: victorspruela@ufmg.br

Resumo—Este trabalho tem como objetivo apresentar uma revisão da literatura de redes neurais artificiais, com enfoque na evoluções das principais técnicas clássicas.

I. INTRODUÇÃO

Redes neurais artificiais (RNA) é uma classe de modelos muito popular em problemas de classificação, reconhecimento de padrões, regressão e predição, sendo aplicado em diversas disciplinas.

A. Problemas de classificação

A tarefa de classificação consiste em associar um conjunto de padrões de entrada, representado por um vetor de características, para uma de varias classes previamente definidas.

B. Problemas de regressão

Dado um conjunto de N pares de dados de entrada-saída $\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)\}$, o objetivo da regressão é encontrar uma função aproximada $\hat{f}(\mathbf{x})$ que melhor descreve a função desconhecida $f(\mathbf{x})$ utilizada para gerar estes dados.

C. Problemas de predição

Considerando um conjunto de N $\{y(t_1), \dots, y(t_N)\}$ amostras ordenadas em função do instante de tempo em que foram amostradas t_1, \dots, t_N , o objetivo da predição é estimar qual será o valor da amostra y_{N+1} em um tempo futuro t_{N+1} .

D. Problemas de reconhecimento de padrões

II. APRENDIZADO SUPERVISIONADO

A. Perceptron simples

B. Máquinas de aprendizado extremo

C. Redes RBF

D. Perceptron de múltiplas camadas

III. APRENDIZADO NÃO-SUPERVISIONADO

A. Aprendizado Hebbiano

B. SOM

IV. GENERALIZAÇÃO

Uma das suas principais características do modelo RNA é sua capacidade de generalização. Em geral, algoritmos de aprendizado supervisionado possuem como objetivo minimizar

o erro quadrático dos valores previstos pelo modelo em relação às saídas em estudo:

$$\sum_{i=1}^N [y_i - f(\mathbf{x}_i)]^2 \quad (1)$$

onde y_i é uma resposta desejada para uma entrada \mathbf{x}_i , e f é o função que aproxima a resposta desejada. Ou seja, estamos interessados em encontrar o conjunto de pesos \mathbf{w} da rede a partir dos pares de dados de entrada-saída $\mathcal{D} = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)\}$ que melhor aproxima a função desconhecida f .

Entretanto, se os dados a serem modelados são ruidosos o uso deste único objetivo pode levar a um overfitting sobre o conjunto de dados de treinamento, de forma que este não consiga generalizar bem para novos valores observados. Estatisticamente, podemos definir a efetividade de f como um estimador de y como [1]:

$$E[(y - f(\mathbf{x}; \mathcal{D}))^2 | \mathbf{x}, \mathcal{D}] = E[(y - E[y|\mathbf{x}])^2 | \mathbf{x}, \mathcal{D}] + (f(\mathbf{x}; \mathcal{D}) - E[y|\mathbf{x}])^2 \quad (2)$$

É importante notar neste indicador que o primeiro termo representa a variância de y dado \mathbf{x} , não dependendo dos dados. Já o segundo termo mede a distância entre o estimador e a regressão. Logo, podemos definir o erro quadrático médio de f como um estimador da regressão $E[y|\mathbf{x}]$ para um conjunto de dados \mathcal{D} como:

$$E_{\mathcal{D}}[(f(\mathbf{x}; \mathcal{D}) - E[y|\mathbf{x}])^2] = (E_{\mathcal{D}}[f(\mathbf{x}; \mathcal{D})] - E[y|\mathbf{x}])^2 \quad \text{“viés”} \\ + E_{\mathcal{D}}[f(\mathbf{x}; \mathcal{D}) - E_{\mathcal{D}}[f(\mathbf{x}; \mathcal{D})]]^2 \quad \text{“variância”} \quad (3)$$

A derivação completa da relação acima pode ser encontrada em [1]. Logo é fácil notar que o aprendizado de RNAs é um problema multi-objetivo, no qual precisamos encontrar uma solução de compromisso entre o viés e a variância do modelo. Portanto, em um dos extremos teremos um conjunto de pesos que resultam em um viés máximo (*underfitting*) e no outro variância máxima (*overfitting*).

A. Máquinas de Vetores Suporte

B. Aprendizado Multiobjetivo

REFERÊNCIAS

- [1] Stuart Geman, Elie Bienenstock, and René Doursat. Neural networks and the bias/variance dilemma. *Neural computation*, 4(1):1–58, 1992.

- [2] Oludare Isaac Abiodun, Aman Jantan, Abiodun Esther Omolara, Kemi Victoria Dada, Nachaat Abdelatif Mohamed, and Humaira Arshad. State-of-the-art in artificial neural network applications: A survey. *Heliyon*, 4(11):e00938, 2018.
- [3] Anil K Jain, Jianchang Mao, and K Moidin Mohiuddin. Artificial neural networks: A tutorial. *Computer*, 29(3):31–44, 1996.