

Utilização de modelos de Perceptron para a classificação de doenças cardíacas

Use of Perceptron models for the classification of heart diseases

RESUMO

Jeferson Felipe Rio Vicente
jefersonrio28@gmail.com
Universidade Tecnológica Federal do Paraná – UTFPR
Apucarana, Paraná, Brasil

Rodrigo dos Santos Veloso Martins
rodrigomartins@utfpr.edu.br
Universidade Tecnológica Federal do Paraná – UTFPR
Apucarana, Paraná, Brasil

As técnicas de aprendizado de máquina realizam a função de simular a tomada de decisões que uma pessoa realiza e podem ser utilizadas em diversas áreas como saúde, economia e indústria. Essas técnicas podem ser utilizadas na avaliação de dados volumosos, acelerando diversos processos. Utilizando redes neurais implementadas no software Scilab, este projeto teve como objetivo classificar corretamente a presença ou ausência de doença cardíaca em pacientes que constam na base de dados “*Cleveland dataset*” presente na página “*Heart disease databases*” e comparar dois modelos de perceptron. Os dois modelos utilizados foram: o perceptron de camada única e o perceptron de múltiplas camadas com algoritmo de retropropagação. A partir desses modelos constatou-se que o modelo de múltiplas camadas obteve maior acurácia, sensibilidade e especificidade, porém com um custo computacional mais elevado

PALAVRAS-CHAVE: Redes Neurais (computação). Aprendizado do computador. Coração – doenças.

Recebido: 19 ago. 2019.

Aprovado: 01 out. 2019.

Direito autoral: Este trabalho está licenciado sob os termos da Licença Creative Commons-Atribuição 4.0 Internacional.



ABSTRACT

Machine learning techniques are capable of simulating the decision making that a person makes and can be used in various areas such as health sciences, economics and industry. These techniques can be used to analyse large volumes data, speeding up various processes. Using neural networks implemented in the software Scilab, this project aimed to correctly classify the presence or absence of heart disease in patients in the “*Cleveland dataset*” database on the “*Heart disease problem*” page and compare two perceptron models. The two models used were the following: single layer perceptron and multilayer perceptron with backpropagation algorithm. From these models it was found that the multilayer model obtained greater accuracy, sensitivity and specificity, but with a higher computational cost.

KEYWORDS: Neural networks (Computer Science). Machine learning. Heart diseases.

INTRODUÇÃO

A Revolução Industrial que ocorreu no século XVIII na Inglaterra foi a precursora de uma nova fase do desenvolvimento tecnológico, onde observou-se a substituição da produção manual pela produção feita por maquinário (HOBBSAWM, 2000). No final do século XX viu-se que o complexo funcionamento de um computador poderia simular o funcionamento de um cérebro humano. Isto gerou novos estudos e possibilidades para o tema. Esse tema foi classificado como uma vertente da ciência da computação, desenvolvendo um estudo chamado de aprendizado de máquinas (HAYKIN, 2007).

O aprendizado de máquinas alinhou conceitos matemáticos e computacionais em um conceito chamado de redes neurais. As redes neurais podem gerar respostas com base nos estímulos a ela atribuída. As redes podem ser utilizadas na previsão de comportamentos e na classificação de padrões, como a identificação de duas raças de cachorros a partir de imagens.

Um dos modelos de redes neurais é o perceptron de Rosenblatt, que consiste em uma rede neural de um único neurônio. Esse é um modelo de camada única e possui apenas um neurônio, que associa os estímulos (dados de entrada) classificando-os, após passarem por uma função de ativação. Um outro modelo de redes neurais é o perceptron de múltiplas camadas. Diferente do de camada única, as informações são processadas mais de uma vez, ou seja, este modelo possui mais de uma camada, cada uma com um neurônio ou mais. O objetivo desse trabalho é avaliar e comparar o perceptron de Rosenblatt e o de múltiplas camadas em um problema de classificação da área da saúde.

MATERIAL E MÉTODOS

Neste estudo os dados obtidos de “*Heart Disease Databases*” (AHA, 2001) foram utilizados para classificar pacientes com presença ou ausência de doenças cardíacas. Nessa base de dados foram computados dados de 297 pacientes de um total de 303 pacientes; os seis pacientes excluídos possuíam dados inconsistentes. As informações presentes de cada paciente incluíam: idade, gênero, tipo de dor no peito, pressão arterial em repouso, colesterol sérico, concentração de açúcar no sangue em jejum, resultados eletrocardiográficos em repouso, ritmo cardíaco máximo alcançado, inclinação do pico de segmento ST durante o exercício, número de grandes vasos coloridos por fluoroscopia talassemia e o diagnóstico do paciente. Os dados representam as 14 coordenadas do vetor de dados.

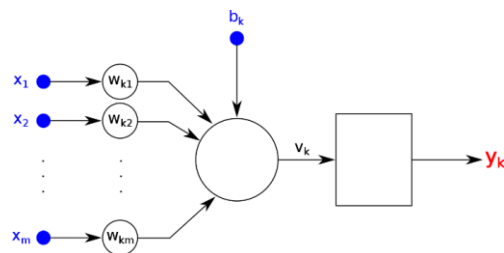
O primeiro algoritmo implementado foi baseado no estudo do perceptron de Rosenblatt de camada única (HAYKIN, 2007). Este é um modelo com um único neurônio, que realiza um produto escalar entre os pesos sinápticos (parâmetros livres) e o vetor de dados inseridos. O resultado deste produto é encaminhado para

a função de ativação que por sua vez atribui a este sinal de entrada a classificação da rede. A função de ativação utilizada nesse caso segue o modelo da Eq. (1).

$$sinal(v) = \begin{cases} +1 & \text{se } v \geq 0, \\ -1 & \text{se } v < 0, \end{cases} \quad (1)$$

onde v representa o produto escalar entre os pesos sinápticos e o vetor de dados e as atribuições +1 e -1 simbolizam respectivamente pacientes com e sem doença cardíaca. Veja a Figura 1. Nessa figura estão representados os dados de entradas x , os pesos sinápticos da rede w , o bias b , o valor v obtido pelo produto entre x e w e pôr fim a resposta y da rede.

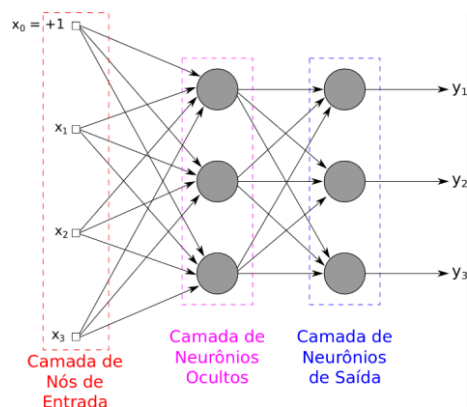
Figura 1 – Modelo de perceptron de camada única



Fonte: Autoria própria.

O outro modelo de perceptron implementado foi o de múltiplas camadas. Este modelo de rede neural possui neurônios como o do perceptron de Rosenblatt como unidades de processamento. Ele realiza o procedimento do produto escalar diversas vezes, de acordo com o número de neurônios em cada camada e o número de camadas (Figura 2). Neste modelo uma camada com M neurônios possui M vetores de pesos sinápticos. Para esse trabalho, o perceptron de múltiplas camadas foi elaborado utilizando-se duas camadas de neurônios, sob o algoritmo de retropropagação de erro (Haykin, 2007). Nessa figura, x representa o dado de entrada, enquanto y representa a resposta da rede.

Figura 2 – Modelo de perceptron de múltiplas camadas



Fonte: Autoria própria.

O algoritmo do perceptron de múltiplas camadas pode ser descrito da seguinte maneira: o algoritmo realiza os produtos escalares entre o sinal de entrada da rede e os vetores de pesos sinápticos dos neurônios da primeira camada. Em seguida as respostas produzidas por essa camada são encaminhadas

para a camada seguinte, sendo processadas de forma semelhante. Este processo se repete sucessivamente até a camada de saída, onde a resposta de um neurônio é obtida a partir da função de ativação escolhida: $\varphi(v) = 1/(1 + e^{-v})$, onde v representa o valor obtido no produto entre os pesos sinápticos e o vetor de dados. Nesse caso, o intervalo de resposta para essa função de ativação é (0,1) e foi atribuída a classificação de presença de doença cardíaca para resposta maior ou igual 0,5 e ausência de doença cardíaca para resposta menor que 0,5.

A execução do algoritmo iniciou-se com o tratamento dos dados, onde foram calculados a média e o desvio padrão de cada variável do problema de modo que sua média fosse 0 e desvio padrão 1. Além disso, houve a separação dos dados em conjuntos de treinamento, desenvolvimento e teste para o perceptron de múltiplas camadas, e conjuntos de treinamento e teste para o perceptron de Rosenblatt. Esta divisão foi feita de modo que a proporção de pacientes com ou sem doença cardíaca fosse igual em todas as etapas do algoritmo.

Os dois modelos considerados passaram por uma fase de treinamento, onde utilizou-se 50% dos dados para o perceptron de múltiplas camadas e 70% para o de de camada única. O vetor de dados de cada paciente do conjunto de treinamento foi apresentado à rede e foram comparadas as repostas produzidas pela rede e a resposta correta. Com essa comparação os pesos sinápticos da rede sofreram correção em caso de resposta incorreta.

No caso do perceptron de camada única, a correção de seu neurônio baseia-se em seu erro, nos dados apresentados a ele e a taxa de aprendizagem, que indicará a intensidade da correção. No perceptron de múltiplas camadas a correção foi feita pelo algoritmo de retropropagação, ou seja, a partir da resposta da camada de saída a rede é corrigida até chegar na primeira camada. Essa retropropagação inicia-se calculando o erro obtido na resposta e o gradiente local, que associa os dados inseridos naquela camada de neurônios, com isso obtém-se um fator de correção para cada peso sináptico. De forma semelhante, a correção da camada oculta acontece utilizando-se o gradiente local de cada neurônio e os pesos sinápticos do mesmo, obtendo assim o fator de correção para esta camada.

O conjunto de treinamento foi percorrido mais de uma vez em ambos modelos, de forma que o perceptron pudesse adaptar-se ao decorrer das iterações. Os dados foram permutados em cada uma das iterações para prevenir um certo vício nas respostas. Ainda na fase de treinamento foram testados dois parâmetros que interferem na convergência da rede: a taxa de aprendizagem (fator que controla a intensidade da correção dos pesos sinápticos) e o número de iterações. Esses testes visavam encontrar em quais condições o menor erro era obtido. Enquanto a fase de treinamento ocorria desenvolveu-se o cálculo de erro pelo método do erro quadrático médio (EQM) a fim de se visualizar o caminho percorrido pela rede e como o erro dela comportou-se durante sua execução.

Após o treinamento, no algoritmo referente ao perceptron de múltiplas camadas, a rede passou pela fase de desenvolvimento, onde os pesos sinápticos obtidos na fase de treinamento foram testados em um novo conjunto de dados. Para essa etapa foram utilizados 30% dos dados da base de dados. Nessa fase, além da visualização da acurácia da rede após o treinamento, foi realizada a verificação da melhor arquitetura que poderia ser atribuída a rede, ou seja, o número de neurônios na camada oculta. As arquiteturas testadas foram executadas com pesos sinápticos iniciais aleatórios no intervalo (-0.5,0.5).

Depois das etapas de treinamento e desenvolvimento, a rede passou pelo que chamamos de fase de testes com os dados restantes, que correspondia a 20% do total no caso do perceptron de múltiplas camadas e 30% no caso do perceptron de camada única. Assim como na fase de desenvolvimento, os dados foram apresentados uma única vez, diferente da fase de treinamento. Essa fase foi realizada separadamente para se avaliar seus resultados. No caso do perceptron de múltiplas camadas, essa fase foi realizada somente com o número de neurônios na camada oculta que forneceu o melhor resultado na fase de desenvolvimento.

Após a realização do mesmo processo para os dois modelos de perceptron, os dados foram avaliados e comparados, para determinar os pontos fortes e fracos de cada modelo e as principais diferenças entre eles.

RESULTADOS E DISCUSSÃO

Para a comparação dos dois modelos, verificou-se a maior acurácia obtida nas execuções da fase de teste, pois utilizou-se alguns parâmetros aleatórios de forma que as execuções dificilmente repetiriam resultados. É importante notar que na fase de teste foram apresentados aos modelos dados nunca antes vistos pelos classificadores. Para o perceptron de camada única, observou-se uma acurácia de aproximadamente 72,2%, já no perceptron de múltiplas camadas pode-se observar uma acurácia de aproximadamente 80%. O resultado do perceptron de múltiplas camadas, melhor dentre os dois, pode ser considerado satisfatório de acordo com trabalhos semelhantes da literatura, que variam entre 77% e 91% nos trabalhos de: Baroni e Da Silva (2018); Rodrigues, Mancrini e Monteiro (2008); Azzi, Do Vale e Da Silva (2018).

No perceptron de múltiplas camadas determinou-se que o melhor resultado foi obtido utilizando-se uma arquitetura possuindo 15 neurônios na camada oculta, utilizando-se 100 iterações no processo de treinamento e com taxa de aprendizagem (η) variando de acordo com a expressão $\eta = 2/(1 + n/30)$. No caso do perceptron de Rosenblatt o valor da taxa de aprendizagem foi decrementado de acordo com $\eta = 0.005 * 0.85^n$.

Uma diferença substancial entre os dois métodos foi a velocidade de execução, onde o perceptron de camada única por ser mais simples realizou suas tarefas muito mais rapidamente que o segundo perceptron. Porém, ao ganhar em velocidade, o perceptron de camada única deixou a desejar em sua acurácia.

Como medida de precisão para os métodos visualizou-se também a sensibilidade e a especificidade das redes. A sensibilidade corresponde à capacidade da rede fornecer um diagnóstico positivo dentre os pacientes com presença de doença cardíaca. A especificidade, por sua vez, expressa a probabilidade de fornecer um diagnóstico negativo em indivíduos sem a doença (BORGES, 2016). Para o modelo de múltiplas camadas a sensibilidade foi de 71,4% e a especificidade foi de 87,5%, no modelo de camada única esses valores foram: 55,8% de sensibilidade e 87,2% de especificidade. Outro ponto importante foi medida da probabilidade da rede indicar falsos negativos, no perceptron de camada única essa probabilidade foi de 21,1% e no modelo de múltiplas camadas foi de 13,3%.

CONCLUSÃO

Os resultados evidenciaram que para esse problema de classificação o modelo de múltiplas camadas obteve melhor resultado, porém com tempo de execução mais lento para o treinamento. Já o perceptron de camada única apresentou grande velocidade, mas com resultados inferiores.

Em relação a literatura, esse modelo de redes neurais apresentou acurácia comparável com outros trabalhos publicados. Como trabalhos futuros pode-se considerar a implementação de técnicas de regularização, de modo a incrementar a capacidade de generalização dos modelos implementados. Acredita-se que esse modelo de algoritmo possa ser empregado no auxílio de diagnósticos por ter boa acurácia e por ter baixo custo operacional, além de fornecer um diagnóstico rápido uma vez treinado.

REFERÊNCIAS

AHA, D. W. Heart Disease Databases. **UCI Machine Learning**. Disponível em: <http://www.ics.uci.edu/pub/machine-learning-databases/heart-disease.names>. Acesso em: 15 mar. 2019.

HAYKIN, S. **Redes Neurais: princípios e práticas**. 2. ed. Porto Alegre: Bookman, 2007.

HOBSBAWN, E. J. **Da Revolução Industrial Inglesa ao Imperialismo**. Rio de Janeiro: Forense Universitária, 2000.

BARONI, Juliana; DA SILVA, Robson Mariano. Utilização de Técnicas de Inteligência Computacional na Caracterização de Pacientes com Doenças Cardiovasculares. **Revista Cereus**, v. 10, n. 4, p. 160-171, 2018.

RODRIGUES, Thiago Baptista; MACRINI, José Leonardo Ribeiro; MONTEIRO, Elisabeth Costa. Seleção de variáveis e classificação de padrões por redes neurais como auxílio ao diagnóstico de cardiopatia isquêmica. **Pesquisa Operacional**, v. 28, n. 2, p. 285-302, 2008.

AZZI, Juliana B.; DO VALE, Italo G.; DA SILVA, Robson M. Utilização de máquina de vetor de suporte na caracterização de pacientes portadores de doenças cardiovasculares. **Proceeding Series of the Brazilian Society of Computational and Applied Mathematics**, v. 6, n. 1, 2018.

BORGES, Leonardo Silva Roeber. Medidas de Acurácia diagnóstica na pesquisa cardiovascular. **Int J Cardiovasc Sci**, v. 29, n. 3, p. 218-22, 2016.