

Utilização de Redes Neurais Artificiais para detecção de Arritmias Cardíacas

Lucas F. Cardoso, Náthalee C. A. Lima

Universidade Federal Rural do Semi-Árido (UFERSA) – Pau dos Ferros/RN – Brasil

lucas.cardosoec@hotmail.com, nathalee.almeida@ufersa.edu.br

Abstract. *This paper presents a brief discussion of neural networks, more specifically about multi-layered Perceptron as a pattern classifier. It is approached in the work the measurement of the use of Software in the medical diagnosis in order to identify pathological as arrhythmia, that is a disease that produces millions of Brazilians and one of the causes of death that has been gaining prominence by its number, being that often such You can not stop using your problem. Learn more about problems and problems with the neural network or diagnosed.*

Resumo. *Este artigo apresenta uma breve discussão acerca de redes neurais, mais especificamente sobre Perceptron de Múltiplas Camadas como um agente classificador de padrões. Aborda-se neste trabalho a importância da utilização de Softwares no diagnostico medico de forma a identificar patológicas como arritmia, que é uma doença que atinge milhões de brasileiros e umas das causas de morte que vem ganhando destaque pelo seu número, sendo que muitas vezes tal paciente nem sabe de determinado problemas e que com a rede neural o diagnostico pode ser feito mais rapidamente possibilitando de forma mais rápida um tratamento.*

Palavras-Chaves: *Arritmia, MLP, Rede Neural*

1. Introdução

Doença cardiovascular, como arritmia, é um termo genérico usado para descrever distúrbios que afetam o coração ou os vasos sanguíneos, sendo hoje uma das principais causas de óbitos em países desenvolvidos e em desenvolvimento JUNIOR (2008). Segundo a Agencia Brasil de Comunicação (EBC), mais 20 milhões de brasileiros sofrem de algum tipo de arritmia cardíaca, doença responsável por mais de 320 mil mortes súbitas no país todos os anos, de acordo com os dados da Sociedade Brasileira de Arritmias Cardíacas (Sobrac). Só no Brasil, a doença arterial coronariana, uma entre os diversos tipos de arritmias, constitui altos índices de internações, estando como uma das maiores causas de mortalidade nos países em desenvolvimento GUS (2015).

Tais índices poderiam ser facilmente reduzidos se previamente fossem diagnosticados. O problema em muitos casos é a quantidade de dados a serem analisados, como também a complexidade muito elevada, impossível de se resolver mentalmente necessitando de uma ferramentas robusta que possa lidar com tal classificação de dados, como é o caso das Redes Neurais Artificiais (RNAs). As RNAs são modelos computacionais inspirados na forma como o sistema nervoso biológico, como o cérebro processa informações (NAHAR, 2012). Sua principal característica tem por base a adaptação e a representação de conhecimentos baseada em conexões.

Dentre as diversas aplicações das RNAs, podemos citar sua utilização para melhoria na tomada de decisões no diagnóstico médico, como exposto por SANTOS (2005) e AMBRÓSIO (2002). Em seu aprendizado, a rede neural é submetida a uma série de diagnósticos de pacientes, de várias características, com vários sintomas e aos resultados de seus testes. Também serão fornecidos os diagnósticos médicos para cada doença. Então, ao serem apresentados novos dados de um novo paciente, com seus sintomas, a rede fornecerá um diagnóstico inicial para os novos casos objetivando uma ferramenta de auxílio ao médico, e não um programa que o substitua.

Com esse intuito, objetivou-se implementar uma rede Perceptron de Múltiplas Camadas (MLP), a fim de distinguir a presença ou ausência de arritmia cardíaca em um grupo de pacientes, dado um conjunto de diagnósticos realizados com cerca de 88 pacientes. Para isso, foi necessário estudar a rede MLP bem como suas principais características além do banco de dados a ser analisado.

Este artigo está organizado da seguinte forma: Na Seção 2, é apresentada a rede MLP e suas principais características, bem como a descrição do conjunto de dados. Na seção 3, serão apresentados os possíveis resultados desta análise e finalmente, na seção 4 são apresentadas as conclusões do trabalho.

2. Desenvolvimento

2.1. Perceptron de Múltiplas Camadas (MLP).

A rede MLP (Perceptron de Múltiplas Camadas), como mostrado na Figura 1, é uma rede que se caracteriza pela presença de no mínimo uma camada neural oculta, localizada entre a camada de entrada e a camada de saída da rede neural. Considerada uma das mais eficientes devido a sua vasta aplicação como, aproximação de funções, reconhecimento de padrões, controle de processos, previsão de séries temporais e otimização de sistemas SILVA e SPATTI e FLAUZINO, (2010).

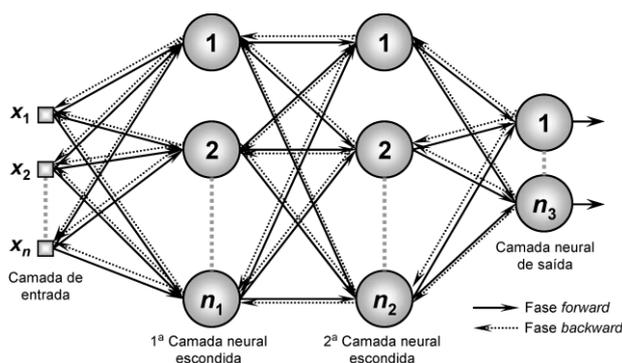


Figura 1: Perceptron de Múltiplas Camadas (MLP)
Fonte: (SILVA e SPATTI e FLAUZINO)

Seu funcionamento consiste em propagar os sinais advindos do meio externo que servem de entrada para a primeira camada de entrada da rede neural. Em cada neurônio da primeira camada, tais sinais são ponderados pelos seus respectivos pesos e limiares produzindo um potencial de ativação que será submetido a sua função de ativação gerando assim um sinal de saída da primeira camada que será usado como entrada para segunda camada oculta. Dessa forma, o sinal de saída dos neurônios de uma camada servem como sinais de entrada para todos os neurônios das camadas

subsequentes, até que a rede produza uma saída final, referente a saída da rede SILVA e PATTI e FLAUZINO (2010).

Tal saída é comparada com a saída desejada, obtendo-se assim uma função de erro, que será aplicada para o ajustes dos pesos, por um algoritmo de treinamento conhecido como *backpropagation*.

2.2. Descrição do conjunto de dados

Como forma de auxiliar no diagnóstico, dentre os problemas de classificação, escolheu-se tratar com arritmias cardíacas, definida como o distúrbio do ritmo cardíaco na qual é estabelecido como normal entre 60 e 100 batimentos por minuto (bpm), HERNANDEZ (2017). Para realizar tal procedimento, foi necessário ter conhecimento de uma base de dados, neste caso, utilizou-se dados do *UCI Machine Learning Repository*.

O banco de dados, que tem por título “*Cardiac Arrhythmia Database*” ou Base de dados de arritmia cardíaca, foi desenvolvido por H. Altay Guvenir, Burak Acar e Haldun Muderrisoglu, em 1998, e tem por base de dados o estudo com 452 casos, cada uma 279 atributos, dentre ele idade, sexo, peso, intervalo da Onda P e T dentre outros, na qual são apresentados diferentes tipos de arritmias, onde pode-se classificar tais casos em 16 grupos, conforme mostrado na Tabela 1.

Tabela 1: Banco de dados Arritmia

Código da classe	Classe	Número de casos
01	Normal	245
02	Alterações isquêmicas (Doença da Artéria Coronariana)	44
03	Velho Infarto Anterior do Miocárdio	15
04	Velho Infarto inferior do Miocárdio	15
05	Taquicardia Sinusal	13
06	Bradicardia Sinusal	25
07	Contração Prematural Ventricular (PVC)	3
08	Contração Prematural Supraventricular	2
09	Bloque de ramo do feixe esquerdo	9
10	Bloque do ramo do feixe direito	50
11	1. Grau AtrioVentricular	0
12	2. Grau Bloco AV	0
13	3. Grau Bloco AV de grau	0
14	Hipertrofia do ventrículo esquerdo	4
15	Fibrilação Atrial ou Flutter	5
16	Outros	22

Dos 16 grupos apresentados acima, apenas dois foram escolhidos para aplicação da rede neural. A Classe 01, considerado normal, ou seja, sem nenhum tipo de arritmia, onde foram escolhidas 43 instâncias. O segundo grupo a ser escolhido foi a classe 02, com as 41 instâncias, que se refere a indivíduos com arritmia cardíaca, mais especificamente alterações isquêmicas (Doença da Artéria Coronariana) que foi

escolhida por ser umas das doenças mais comuns e a primeira causa de óbito nas sociedades modernas segundo SICARI (2008).

2.3. Avaliação do Conjunto de dados

Ao desenvolvermos uma aplicação utilizando redes neurais artificiais, são necessários antes, uma série de procedimentos. Inicialmente os dados utilizados para classificação foram minuciosamente coletados, embaralhados e devidamente separados em dois conjuntos: o conjunto de treinamento que possuem 40 instâncias e um conjunto de validação que possuem 44 instâncias, ambas possuindo dados das classes 01 e 02.

No conjunto de treinamento, os dados são apresentados bem como suas respectivas saídas, de forma que ocorra um ajuste nos pesos sinápticos, que serão utilizados para validação da rede. Já no conjunto de teste, novos dados são apresentados a rede e com os pesos sinápticos já ajustados no treinamento, a rede fará o papel de classificar de forma mais precisa possível, de acordo com a eficiência do aprendizado.

Ao serem separados em grupos de validação e treinamento, um conjunto de parâmetros que irão compor as configurações da rede neural são escolhidos, sejam eles o erro mínimo do conjunto de treinamento ou a quantidade de épocas, a quantidade de camadas, a quantidade de neurônios por camada entre outros que vão variar de acordo com o problema. Por último, vendo o desempenho da rede neural, a mesma pode ser posteriormente implementada.

Um das formas de avaliar o conjunto de dados é utilizando uma matriz de classificação, também conhecida como matriz de confusão. Essa matriz é uma ferramenta importante que determina se o valor previsto (classificado) correspondeu ou não ao valor real, de forma que todos os casos em cada categoria são contabilizados e os totais são exibidos na matriz. Uma representação detalhada é mostrada na Figura 2.

		Valor Verdadeiro	
		Positivo	Negativo
Valor Previsto	Positivo	Verdadeiros Positivos	Falsos Positivos
	Negativo	Falsos Negativos	Verdadeiros Negativos

Figura 2: Matriz de confusão

Conforme ilustrado na Figura 2, as linhas na matriz representam os valores previstos para o modelo, e as colunas representam os valores atuais ou verdadeiros, onde na diagonal principal são representados os resultados corretos e na diagonal secundária os valores previstos erroneamente. Valores “Verdadeiros Positivos” e “Verdadeiros Negativos” são valores que obtiveram valores previstos iguais aos valores verdadeiros, ou seja, foram previstos como positivos e negativos corretamente, de acordo com os dados reais. Os valores “Falsos Positivos e Falsos Negativos” são valores

errôneos, que respectivamente, foram previstos como positivos, quando seu dado real era negativo ou foram previstos como negativos quando seu dado real era positivo.

2.4. Configuração do *Perceptron* Múltiplas Camadas

Para implementar a rede MLP foi utilizado o *Software Matlab* que além de robusto, disponibiliza um conjunto de funções para utilização em RNAs.

Os parâmetros selecionados para implementar a rede foram escolhidos de forma a se obter o menor erro possível entre as saídas da rede e as saídas desejadas. Tais parâmetros são mostrados na Tabela 2, mostrada abaixo.

Tabela 2: Parâmetros MLP

Número de camadas	3
Números de camadas escondidas	2
Neurônios na primeira camada escondida	10
Neurônios na segunda camada escondida	10
Número de neurônios na camada de saída	1
Número de dados de entrada	279
Função de ativação	Tangentes hiperbólica
Erro	10^{-20}
Algoritmo de treinamento	Levenberg-Marquardt

A ilustração da rede MLP, é ilustrada na Figura 3, mostrada abaixo.

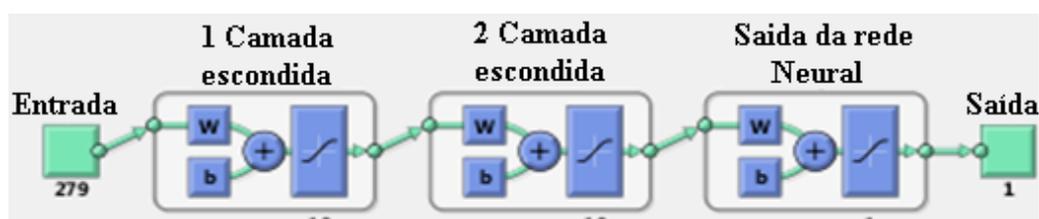


Figura 3: Implementação MLP

Para representar as duas classes foram utilizados valores como -1 para classificar pessoas da classe 01, ou normais e +1 para pessoas que apresentam arritmia cardíaca. Neste caso, pela saída apresentar um range de valores entre -1 e 1, escolheu-se a função de transferência para cada neurônio como tangente hiperbólica devido suas características serem mais suavizadas e possibilitar tais resultados.

3. Resultados e Discussões

Assim, para os nove treinamentos realizados, foram aplicados ao MLP cujos resultados foram registrados na Tabela 3. As saídas representados por 1 são referentes a pessoas classificadas com Arritmia, já as saídas classificadas com -1, são pessoas consideradas normais.

Como forma de representar se a rede teve uma bom desempenho, incluiu-se na Tabela 3, uma coluna com resultado final, que tem por objetivo nos mostrar se a moda dos dados, ou seja, o valor que ocorre com maior frequência, foram iguais as saídas esperadas. Em caso afirmativo, temos como resultado, Y de (SIM), e em caso negativo, N, de (Não), conforme segue mostrado abaixo.

Tabela 3: Dados de saída da Rede MLP

Saídas esperadas	Dados de Saída da Rede MLP									Resultado Final
1	1	1	1	1	1	1	1	1	1	y
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	y
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	y
-1	-1	-1	-1	-1	1	-1	-1	-1	-1	y
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	y
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	y
1	1	1	1	1	1	1	1	1	1	y
1	-1	-1	-1	-1	-1	-1	1	-1	1	n
1	1	1	1	1	1	1	1	1	1	y
-1	-1	-1	1	-1	-1	-1	-1	-1	-1	y
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	y
1	1	1	1	1	1	1	1	1	1	y
1	1	1	1	1	1	1	1	1	1	y
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	y
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	y
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	y
1	1	1	1	1	1	1	1	1	1	y
1	1	1	-1	1	1	1	1	1	1	y
1	1	1	1	1	-1	1	1	1	1	y
1	1	1	1	1	1	1	1	1	1	y
1	1	1	1	1	1	1	1	1	1	y
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	y
-1	-1	-1	-1	1	-1	1	-1	-1	-1	y
1	1	1	-1	1	-1	1	1	-1	-1	y
1	1	1	1	1	1	1	1	1	1	y
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	y
1	1	1	1	1	1	1	1	1	1	y
1	1	1	1	1	1	1	1	1	1	y
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	y
1	1	1	1	1	1	1	1	1	1	y
1	1	1	1	1	1	1	1	1	1	y
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	y
1	1	1	1	1	1	1	1	1	1	y
1	1	1	1	1	1	1	1	1	1	y
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	y
1	1	1	1	1	1	1	1	1	1	y
-1	1	1	1	1	1	1	1	1	1	n
-1	-1	-1	1	-1	-1	-1	-1	-1	-1	y
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	y
-1	1	-1	1	1	1	-1	1	-1	-1	n
1	-1	-1	-1	1	-1	-1	-1	1	1	n
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	y

1	1	1	1	1	1	-1	1	1	1	y
1	1	1	1	1	1	1	1	1	1	y
-1	-1	-1	-1	-1	1	-1	-1	-1	-1	y
1	1	1	1	1	1	1	1	1	1	y
1	-1	-1	-1	-1	-1	-1	-1	-1	-1	n

Com base nos dados acima, montou-se a matriz de confusão, onde foi levado em consideração a moda dos resultados de saída da MLP, conforme mostra a Tabela 4

Tabela 4: Matriz de confusão

		Valor Verdadeiro		
		<i>Normal</i>	<i>Arritmia</i>	<i>Total</i>
Valores previstos	<i>Normal</i>	20	3	23
	<i>Arritmia</i>	2	19	21
	<i>Total</i>	22	22	

A partir da matriz confusão verificou-se o desempenho da MLP para classificar uma pessoa sem patologia cardíaca (normal) a uma taxa 87%, enquanto para classificar pessoas diagnosticadas com arritmia, obteve-se uma taxa de 90,5%. Dessa forma, calculou-se a eficiência da MLP em 88,6%, calculada a partir da equação abaixo:

$$Eficiencia = \frac{VN}{VN + FP} + \frac{VP}{VP + FN} \quad (1)$$

Onde de acordo com a Figura 2, temos:

VP - VERDADEIRO POSITIVO

FP - FALSO POSITIVO

VN - VERDADEIRO NEGATIVO

FN - FALSO NEGATIVO

4. Conclusão

As redes neurais artificiais utilizadas como um classificador são consideradas ferramentas de grande auxílio no diagnóstico médico e tem se mostrado eficiente, principalmente para diagnosticar doenças, como no caso deste trabalho, na detecção de patologias cardíacas como Arritmia, que vem ganhando destaque ao longo dos anos, devido a grande quantidade de casos.

Com base em *softwares* computacionais e dado o crescente estudo sobre redes neurais, problemas que antes necessitavam de um grande processamento e um vasto tempo de classificação, hoje podem ser representados de forma mais rápida. Dentre as diversas topologias de redes neurais, escolheu-se a rede *Perceptron* de Múltiplas Camadas, com o intuito de classificar dados de pacientes com arritmia. O sistema proposto teve por base a coleta de dados em uma base de dados disponível na internet, utilizando o algoritmo, desenvolvido pelos próprios autores como forma de prover um modelo neural para o diagnóstico de arritmias, que embora tenha se obtido uma taxa de

erro de 11,3%, o resultado se mostrou satisfatório, pois em muitos casos, como no trabalho de SANTOS (2005), utilizou-se um modelo de regressão linear para identificar quais variáveis eram mais relevantes, de forma a se minimizar o erro e que por sua vez, obteve-se um desempenho aproximado em relação a este trabalho.

Referencias

- Arritmias cardíacas causam 320 mil mortes súbitas por ano, alerta entidade. Disponível em: <<http://agenciabrasil.ebc.com.br/geral/noticia/2016-11/arritmias-cardiacas-causam-320-mil-mortes-subitas-por-ano-alerta-entidade>> Acessado em: 28 de Fevereiro de 2018.
- Arrhythmia Data Set. Disponível em: <<https://archive.ics.uci.edu/ml/datasets/arrhythmia>>. Acessado em: 28 de Fevereiro de 2018.
- AMBRÓSIO, Paulo Eduardo. Redes neurais artificiais no apoio ao diagnóstico diferencial de lesões intersticiais pulmonares. 2002. Tese de Doutorado. Universidade de São Paulo.
- Doença arterial coronariana. Arq. Bras. Cardiol., São Paulo , v. 93, n. 6, supl. 3, p. 282-286, Dec. 2009 . Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0066-782X2009001500005&lng=en&nrm=iso>. Acessado em: 28 Fevereiro de 2018.
- Dua, D. and Karra Taniskidou, E. (2017). UCI Machine Learning Repository [<http://archive.ics.uci.edu/ml>]. Irvine, CA: University of California, School of Information and Computer Science.
- GUS, Iseu et al. Variações na prevalência dos fatores de risco para doença arterial coronariana no Rio Grande do Sul: uma análise comparativa entre 2002-2014. Arq Bras Cardiol [Internet], p. 573-9, 2015.
- HERNÁNDEZ, R. Vallejo et al. Arritmias. Panorama actual del medicamento, v. 41, n. 404, p. 524-531, 2017.
- JÚNIOR, SÉRGIO RENATO ROGAL. Detecção e classificação de arritmias cardíacas utilizando redes neurais artificiais auto-organizáveis. Tese de Doutorado. Ph. D. dissertation, Pontifícia Universidade Católica do Paraná. 2008.
- Lichman, M. (2013). UCI Machine Learning Repository [<http://archive.ics.uci.edu/ml>]. Irvine, CA: Universidade da Califórnia, Escola de Informação e Ciência da Computação.
- NAHAR, K. Artificial Neural Network. COMPUSOFT, v. 1, n. 2, p. 25–27, dez. 2012.
- SANTOS, Alcione Miranda dos et al. Usando redes neurais artificiais e regressão logística na predição da hepatite A. Revista Brasileira de Epidemiologia, v. 8, p. 117-126, 2005.
- Sicari R, Nihoyannopoulos P, Evangelista A, Kasprzak J, Lancellotti P, Poldermans D, et al; European Association of Echocardiography. Stress echocardiography expert consensus statement: European Association of Echocardiography (EAE). Eur J Echocardiogr. 2008; 9 (4): 415-37.

SILVA, IN da; SPATTI, Danilo Hernane; FLAUZINO, Rogério Andrade. Redes neurais artificiais para engenharia e ciências aplicadas. São Paulo: Artliber, 2010.