

Classificação de sinais EGG combinando Redes Neurais e Análise de Componentes Independentes

Hallan Santos¹, Carlos A.E. Montesco¹, Methanias C. Júnior¹

¹Universidade Federal de Sergipe – Sergipe, SE – Brasil

{hallancosmo,estombelo}@gmail.com, mjrse@hotmail.com

Abstract. *Identify digestive traits of people through Electrogastrography (EGG) is important because it is usually an economic, non-invasive and less bother than the traditional Endoscopy procedure. This research evaluates the behavior of artificial neural networks learning on the components extracted by Independent Component Analysis (ICA) algorithms. An experiment with statistical analysis whose goal was to present the relationship between the viewing of neutral, negative or positive images and digestive reactions was performed. The results showed that extract only the stomach signal component may reduce the error rate of learning of the neural network compared with experiment.*

Resumo. *Identificar características digestivas de pessoas através da Eletrogastrografia (EGG) é importante pois esta costuma ser uma opção barata, não-invasiva e incomoda menos que o tradicional procedimento de Endoscopia. Este trabalho avalia o comportamento do aprendizado das redes neurais artificiais diante de componentes extraídas por algoritmos de Análise de Componentes Independentes (ICA). Foi realizado um experimento com análise estatística cujo objetivo foi apresentar a relação entre a visualização de imagens neutras, negativas ou positivas e as reações digestivas. Os resultados mostraram que extrair apenas a componente do sinal do estômago deve reduzir a taxa de erro do aprendizado das redes neurais artificiais diante do experimento realizado.*

1. Introdução

Uma das formas de medir as atividades elétricas do estômago é a Eletrogastrografia (EGG). Este procedimento grava os sinais elétricos que passam através dos músculos desta região. São estes sinais que proporcionam contrações diferentes a depender da situação a qual alguém é submetido [Gopu et al. 2009]. A EGG tem se tornado uma ferramenta atraente para profissionais da área de saúde porque, em geral, causa pouco desconforto aos seus pacientes, além de apresentar menos riscos por ser uma técnica não-invasiva. Com o avanço das pesquisas em processamento de sinais, torna-se possível inferir informações importantes [Hubka et al. 2005] e realizar diagnósticos mais precisos a partir dos dados adquiridos através da EGG.

Uma das maiores dificuldades ao se processar os dados obtidos por este procedimento é a captura de outros sinais que não são de interesse, como os sinais respiratórios ou cardíacos. Estes sinais, normalmente possuem maior energia e frequência [Wang 1997], tornando a leitura dos sensores uma mistura entre o sinal de interesse, vale dizer, os sinais da região do estômago, e um conjunto de outros sinais.

Neste cenário, surge a ICA – Análise de Componentes Independentes – como uma alternativa de obter-se o sinal de interesse. O objetivo com esta técnica é separar um conjunto de sinais misturados em um conjunto de sinais independentes [Comm 1994]. Uma vez que os sinais estejam separados, é possível, por exemplo, utilizar apenas aqueles que sejam de interesse para realizar o treinamento de uma rede neural.

Este trabalho realizou um comparativo entre todas as componentes estimadas pela Análise de Componentes Independentes, utilizando três funções distintas. Estas componentes foram processadas por três redes neurais artificiais diferentes, através do algoritmo Backpropagation, da regra delta-bar-delta e do algoritmo de LevenbergMarquardt.

O artigo está organizado da seguinte forma: a seção 2 apresenta os algoritmos e conceitos gerais de ICA aplicados aos propósitos deste trabalho; as redes neurais são introduzidas na seção 3, na qual são apontadas as principais variáveis dos modelos e as otimizações de interesse; outros trabalhos relacionados a esta pesquisa são apresentados na seção 4; a seção 5 exhibe o experimento realizado neste trabalho, levantando objetivos, questões, métricas e hipóteses; os resultados estão apontados na seção 6; as considerações finais, discussão e os trabalhos futuros estão na seção 7.

2. Análise de Componentes Independentes - Algoritmo FastICA

Pode-se dizer que a principal motivação ao estudar a Análise de Componentes Independentes é o problema da separação cega de fontes [Hyvärinen 1999]. No contexto desta pesquisa, ao ser tomada uma Eletrogastrografia com algumas fontes de sinais, por exemplo, regiões diferentes do estômago, existe um sensor para medir especificamente o sinal de cada fonte. Em um cenário ideal, o sinal obtido por cada sensor deveria vir apenas da região específica do estômago, este sinal se denomina sinal de interesse. O que se observa, no entanto, é que os sensores, que deveriam medir apenas os sinais de interesse, na verdade medem uma soma ponderada de sinais provenientes de todas as fontes. Esta mistura de sinais se deve a diversos fatores como a localização do órgão, as propriedades acústicas do próprio aparelho digestivo e os ruídos de outros órgãos do respiratório ou cardíaco, conforme o caso. O desafio em ICA é estimar os sinais originais de cada fonte utilizando apenas os sinais observados pelos sensores [Hyvärinen et al. 2001].

Costuma-se assumir que o processo de mistura dos sinais de interesse se dá através de uma transformação linear, onde uma matriz A pondera todas as fontes, ditas componentes independentes. Em um caso particular, onde apenas dois sinais são as fontes e há dois sensores, o problema da separação cega de fontes pode ser representado da seguinte forma $\mathbf{s} = [s_1 \ s_2]$, $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ e $\mathbf{x} = [x_1 \ x_2]$, onde, $\mathbf{x} = [x_1 \ x_2]$
 $= [s_1 \ s_2] * \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$. A equação ainda pode ser escrita na forma $\mathbf{x} = \mathbf{s} * A$, onde \mathbf{x} é um vetor aleatório cujos elementos representam os sinais observados e a matriz A , cujos valores a_{ij} , linha i e coluna j , representam uma ponderação do vetor aleatório \mathbf{s} , que armazena todos os sinais originais das fontes. Por esta modelagem, deseja-se determinar o vetor aleatório \mathbf{s} , conhecendo-se apenas o vetor aleatório \mathbf{x} [Hyvärinen et al. 2001].

Uma técnica desenvolvida para resolver este problema é a análise de componentes independentes, ou ICA. Ela pode ser utilizada para estimar os valores a_{ij} que compõem a matriz de mistura A baseando-se na informação de independência e **não-gaussianidade das fontes** que formam s em cada instante de tempo [Hyvärinen et al. 2001].

O FastICA é um algoritmo proposto por Hyvärinen com o objetivo de estimar o vetor aleatório s de componentes independentes, através de uma matriz de separação W , inversa da matriz de mistura A . O ideal é que o vetor estimado s seja o mais próximo dos sinais originais quanto possível. Para tanto, as componentes são estimadas utilizando-se um algoritmo de ponto-fixado [Todd 1976], capaz de encontrar, com boa eficiência, um ponto de maximização dos valores de uma função (e^x , por exemplo) aplicada para a estimativa do algoritmo. Um algoritmo de ponto-fixado é um método mais geral que o FastICA, portanto outras técnicas em ICA podem ser implementadas com o modelo de ponto-fixado obtendo assim, considerável ganho de performance. Existe um conjunto de boas funções [Hyvärinen 1997] para execução do FastICA, no entanto, novas funções podem apresentar resultados melhores para conjunto de dados específicos. Para avaliar as diferenças que funções distintas podem produzir, este trabalho utiliza três implementações com funções diversas: **logcosh(x)**, e^x e x^3 .

3. Redes Neurais Artificiais

As redes neurais artificiais têm sido utilizadas em diversos sistemas graças à sua flexibilidade. Suas aplicações encontram-se em diversas áreas, como reconhecimento de caracteres ou previsão de séries temporais, atualmente, já há pesquisas apontando para a análise de imagens médicas com sistemas de diagnóstico [Gopu et al. 2009, Yu et al. 2006, Hubka et al. 2005].

Basicamente, é desejável que uma rede neural aprenda a partir de exemplos, se adapte ao diversos tipos de entrada, tenha capacidade de generalização e seja tolerante a ruídos [Fausset 1994]. A topologia escolhida ao modelar a rede neural é algo crítico pois determina o desempenho da solução, por isso, deve ser tomada empiricamente ou de acordo com a indicação de boas literaturas [Kovacs 1996].

O elemento fundamental da rede neural modelada neste trabalho é o neurônio de McCulloch & Pitts [Haykin 1999]. Estes neurônios são distribuídos em camadas, sendo ao menos uma camada de entrada, uma camada de saída, e pelo menos uma camada intermediária. Este modelo ainda relaciona cada entrada a um peso e fornece uma função de ativação durante o processamento. Uma rede neural multicamada (MLP) pode ser definida como uma rede neural que utiliza mais de uma camada intermediária. Sob a MLP, muitos algoritmos podem ser implementados. O algoritmo *Backpropagation* é um dos mais importantes.

O *Backpropagation* é um algoritmo supervisionado, ou seja, aprende através de exemplos. Apesar de ser muito utilizado, este algoritmo não garante a convergência, isto faz com que dependendo dos valores iniciais de seus parâmetros, a rede possa divergir, ou até mesmo, convergir para um mínimo local. Enquanto o *Backpropagation* utiliza a descida de gradiente como método de aproximação do mínimo de uma função erro, o algoritmo de Levenberg-Marquardt (LM) utiliza uma aproximação pelo método de Newton. Esta abordagem, produz resultados mais precisos que o *Backpropagation* puro.

O motivo é que seu efeito pode ser desfeito nas situações em que o erro aumenta após a atualização dos pesos do *Backpropagation* na fase de *Backward*.

Uma outra otimização para o algoritmo *Backpropagation* é a regra deltabar-delta. Esta regra exige que seja criada uma taxa de aprendizagem para cada peso da arquitetura MLP. Isto faz com que a variável que registra a taxa de aprendizagem seja uma matriz da mesma dimensão que a matriz \mathbf{W} de pesos, aumentando a complexidade da solução. Mais detalhes sobre as técnicas mencionadas encontram-se em [Haykin 1999, Fausset 1994].

4. Trabalhos relacionados

A análise de imagens de EGG utilizando a combinação de ICA e redes neurais artificiais foi introduzida em [Wang 1997] com a justificativa de que o ICA não precisa de informações a priori sobre os sinais das fontes, e isto é o que ocorre em muitas situações práticas da medicina. Wang realizou um experimento com apenas duas funções cosseno e uma de ruído. A sua tentativa mostrou que esta abordagem, além de ser mais simples de implementar, possuía uma rápida convergência.

Em [Hubka et al. 2005], algoritmos de ICA foram implementados no dispositivo *ProGastro*, que realizou a separação das componentes através do método da máxima verossimilhança. O trabalho mostrou que o ICA pode levar a uma ferramenta promissora de separação cega de fontes. Conseguiu-se nesta pesquisa detectar a incidência de contração no estômago a partir da leitura de quatro sinais misturados, contudo, não foi possível obter resultados melhores, pois o dispositivo não lidou bem com a baixa amplitude do sinal de interesse, sendo esta facilmente sobreposta por outros espectros, como o batimento cardíaco.

O algoritmo *Backpropagation* é utilizado em [Yu et al. 2006] para classificação das componentes independentes de eletrocardiogramas (ECG). Neste trabalho, apenas algumas componentes, consideradas importantes, foram selecionadas dentre as componentes estimadas pelo ICA. Ao realizar o treinamento da rede neural, utilizando 23 componentes de entrada e 8 classes, o trabalho conseguiu obter uma precisão de 98.37%, evidenciando o potencial da combinação entre ICA e redes neurais.

[Gopu et al. 2009] propôs uma nova abordagem para a gravação de sinais de EGG que utilizou filtros e uma rede neural para classificação de 180 pacientes sofrendo de desordens no estômago. Este trabalho utilizou métodos sofisticados para o treinamento das redes neurais, como *Learning Vector Quantization* (LVQ) e *Adaptive Resonance Theory* (ART). Após o treinamento, o LVQ e o ART conseguiram obter uma taxa de precisão de 98.57% e 85.71%, respectivamente.

Assim como [Yu et al. 2006] fez para sinais de ECG, este trabalho realizou a classificação de sinais de EGG utilizando redes neurais, diferenciando-se, pois utiliza um algoritmo de maximização de não-gaussianidade, FastICA, considerando todas as suas componentes estimadas para o treinamento da rede neural. Além disto, foi realizada uma comparação entre o algoritmo de Levenberg-Marquadt e a regra deltabar-delta, fazendo uma análise estatística do experimento.

5. Experimento

O processo de experimentação foi categorizado em quatro fases por Víctor Basili: definição, planejamento, operação e interpretação [Basili et al. 1986]. Estas fases foram refinadas e promoveram o *Goal-Question-Metric* (GQM) [Basili et al. 1994], que define um esquema de experimentação em três níveis: o objetivo que se deseja alcançar com o experimento (*goal*), as perguntas que devem ser respondidas para se alcançar os objetivos (*questions*), e as métricas que medem as respostas das perguntas (*metrics*).

Do ponto de vista do GQM, este experimento irá **analisar** as componentes independentes estimadas pelo algoritmo FastICA em uma rede neural **com a finalidade de** avaliar se existe relação entre a visualização de imagens neutras, negativas ou positivas e os sinais observados da região do estômago **com respeito** ao aprendizado da rede neural **do ponto de vista de** pesquisadores e profissionais da área médica **no contexto da** separação cega de sinais e das redes neurais.

5.1. Formulação de hipóteses

A formulação GQM com objetivos, perguntas e métricas encontra-se na tabela 1. Como é necessário garantir a não-gaussianidade dos dados das fontes neste experimento, deve-se utilizar o teste de Kolmogorov-Smirnov (KS) [Galen et al. 1986]. Além disto, o experimento se propõe a comparar o erro quadrático médio (MSE) obtido pelo tradicional algoritmo *Backpropagation*, pelo algoritmo LM e pela regra delta-bar-delta. Assim, foi utilizado o teste não-paramétrico Kruskal-Wallis [Corder et al. 2009]. Podese observar na tabela 2 a formulação destas hipóteses.

Tabela 1 – Formulação GQM do objetivo, perguntas e métricas

<u>Objetivo</u>	<ul style="list-style-type: none">• Avaliar a relação entre as componentes estimadas por ICA e visualização de imagens neutras, negativas e positivas.
<u>Perguntas</u>	<ul style="list-style-type: none">• As funções utilizadas para ICA fazem a rede neural convergir?• Os algoritmos de treinamento e otimização da rede neural impactam nas taxas de acerto.
<u>Métricas</u>	<ul style="list-style-type: none">• Erro quadrático médio (MSE);• Número de iterações para estimar as componentes;• Número de iterações até a convergência dos algoritmos utilizados na rede neural.

Tabela 2 - Formulação das hipóteses

Hipóteses	Teste estatístico
H ₀ : A amostra é gaussiana. H ₁ : A amostra não é normal e pode ser utilizada para o experimento.	KolmogorovSmirnov
H ₀ : Os algoritmos produzem o mesmo MSE. H ₁ : O MSE produzido por algum algoritmo é menor.	Kruskal-Wallis

4.2. Instrumentação

Os dados são coletados em laboratório através de quatro sensores que medem os sinais de regiões específicas do estômago de diversos pacientes. Cada sensor realiza 1,5 milhão de medições por minuto. Em cada período de dois minutos, uma pessoa está observando imagens relacionadas a sentimentos neutros, negativos, e positivos, nesta ordem. Formando assim, três conjuntos de dados, rotulados como neutro, negativo e positivo. Foi utilizada a ferramenta R para execução do algoritmo FastICA, implementação do *Backpropagation* e dos algoritmos de otimização de redes neurais e para utilização das bibliotecas que realizem os testes estatísticos.

6. Resultados

O teste de normalidade Kolmogorov-Smirnov (KS) realizado em R sob todo o conjunto de dados produziu um p-valor médio para os quatro sensores de 10^{-16} . Com isto, **rejeitase a hipótese H_0** levantada na seção 5.1 de que os dados são normais, e portanto, podese admitir que os dados são não-gaussianos.

Ainda sobre o teste de KS, é possível observar a máxima diferença absoluta entre a função de distribuição de fato dos dados, obtida através de algum procedimento empírico, e a função de distribuição acumulada que se poderia assumir dos dados coletados, neste caso a distribuição gaussiana. Como ocorre nos demais testes estatísticos, compara-se esta diferença com um valor crítico, denominado nível de significância. Neste sentido, pode-se notar na figura 1 a distância entre a distribuição empírica dos dados e a gaussiana, reforçando a decisão de rejeitar a hipótese de que os dados são normais.

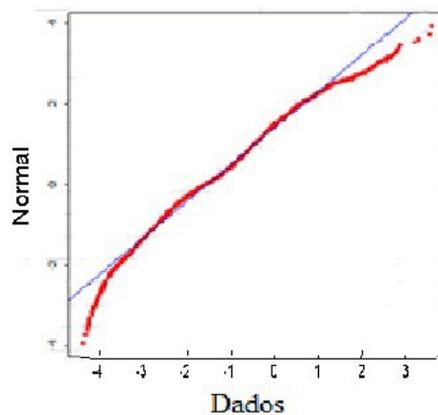


Figura 1. Distribuição do conjunto de dados no teste KS

Para o algoritmo FastICA, que separa as componentes, foram utilizadas as funções: **logcosh(x)**, e^x e x^3 . Em todos os casos, o algoritmo convergiu. Após rodar dez vezes cada um deles, nos três tipos conjuntos de dados abordados (neutro, negativo e positivo), o número médio de iterações foi: 3.83, 4.27 e 2.21, respectivamente.

Cada quádrupla (referente aos quatro sensores) de componentes independentes estimadas pelo algoritmo FastICA foi rotulada como neutra, negativa ou positiva. Analisando a densidade espectral das componentes destes rótulos, pode-se observar a forte presença da frequência 0.05Hz em pelo menos uma das componentes. [Komorowski 2009] apresenta a distribuição destas frequências em função do

comportamento digestivo. Pode-se observar na figura 2, a função densidade espectral destas componentes.

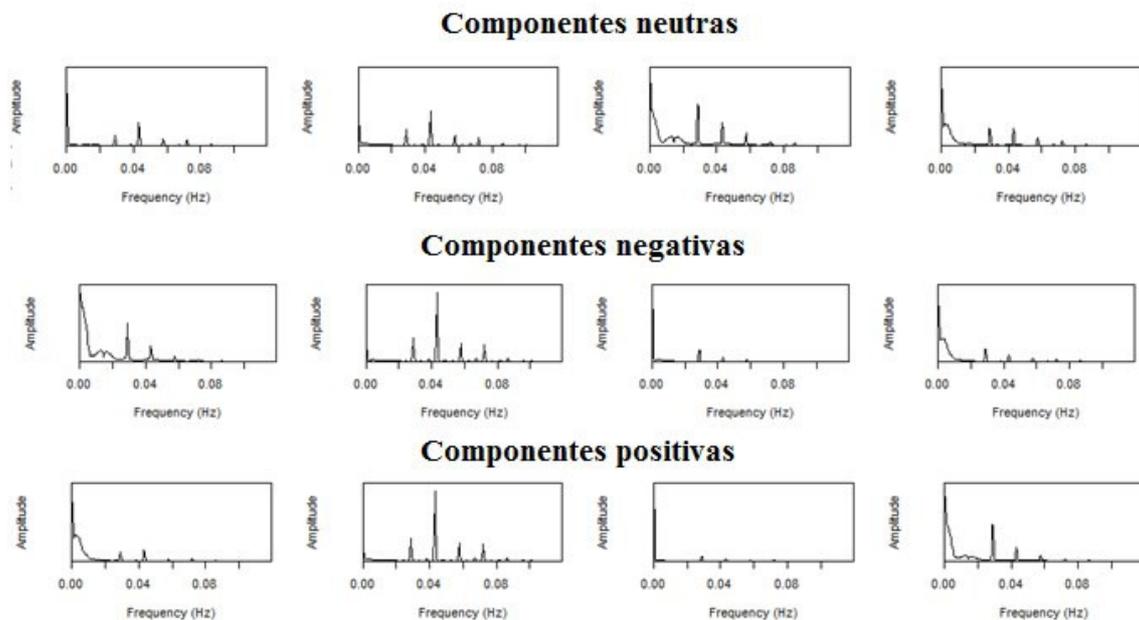


Figura 2 - Densidade espectral das quatro componentes estimadas para cada tipo de conjunto de dados

Este conjunto de quádruplas de componentes independentes estimadas foi dividido de forma que 70% fosse usada para treinamento, 20% para validação cruzada e 10% para testes, conforme orientado em [Haykin 1999], para realização do aprendizado das redes neurais.

Um dos parâmetros ajustados nas redes neurais foi a taxa de aprendizagem (η). Quando esse valor aumenta, houve uma sensível perda de precisão da rede neural. O valor $\eta = 2.5$ em particular produziu algumas vezes resultados piores que $\eta = 1.5$. Isto pode ser explicado graças ao fato de que quando η aumenta seu valor começar a atingir uma zona em que o algoritmo tende a divergir sempre.

Com a utilização da validação cruzada durante a fase de treinamento, a rede neural produziu duas matrizes de pesos, uma associada ao momento em que o treinamento se encerra com o valor mínimo ao ser atingido o limite de iterações, e outra ao instante em que a validação cruzada atinge seu menor valor. Observa-se nos resultados que em 74% das saídas, o MSE é menor quando se utiliza a matriz de pesos obtida no momento em que o erro de validação é mínimo. Nos demais casos, a matriz de pesos possui resultados mais satisfatórios. Este já era o comportamento enunciado em [Haykin 1999]. Segundo o autor, ao não parar no momento em que a validação cruzada atinge seu mínimo, a rede tende a perder generalidade, piorando seus resultados. No entanto, neste trabalho, em 26% dos resultados conseguiu-se um menor MSE utilizando a matriz de pesos referente ao limite de iterações, indicando que o algoritmo pode ter encontrado um novo mínimo local, melhor que o anterior.

Pode-se observar nas tabelas 3, 4 e 5 os melhores resultados dos três algoritmos após realizar uma série de variações nos parâmetros indicados nas tabelas. As redes não convergiram até o limite de iterações estabelecido, que foi 5000. A coluna Min Tre.

indica o MSE ao atingir o limite de iterações, enquanto a coluna Min Val. aponta o MSE quando o erro da validação cruzada foi mínimo.

Em geral, quando $\eta = 0.5$ ou $\eta = 1.5$, o algoritmo LM apresentou discretamente menores erros em relação aos demais algoritmos. Quando $\eta = 2.5$, a regra delta-bardelta obteve um erro um pouco menor. Estas diferenças de desempenho são avaliadas posteriormente através do teste de Kruskal-Wallis e os menores valores estão destacados em vermelho nas tabelas.

Tabela 3. Melhores resultados do algoritmo *Backpropagation*

η	0.5		1.5		2.5	
	Min Tre.	Min Val.	Min Tre.	Min Val.	Min Tre.	Min Val.
MSE	0.6704423	0.6709747	1.042394 7	0.6772700	0.694529 1	0.694947 6

Tabela 4. Melhores resultados da Regra delta-bar-delta para $\eta = 2.5$

β	0.3		0.6		0.6	
K	0.55		0.01		0.55	
γ	0.50		0.033		0.50	
	Min Tre.	Min Val.	Min Tre.	Min Val.	Min Tre.	Min Val.
MSE	0.6765353	0.6704636	0.666216	0.6682878	0.6760326	0.6704447

Tabela 5. Melhores resultados do algoritmo Levenberg-Marquardt

η	0.5		1.5		2.5	
	Min Tre.	Min Val.	Min Tre.	Min Val.	Min Tre.	Min Val.
MSE	0.6700545	0.6684148 7	1.035184 8	0.6118184	0.690122 4	0.688484 1

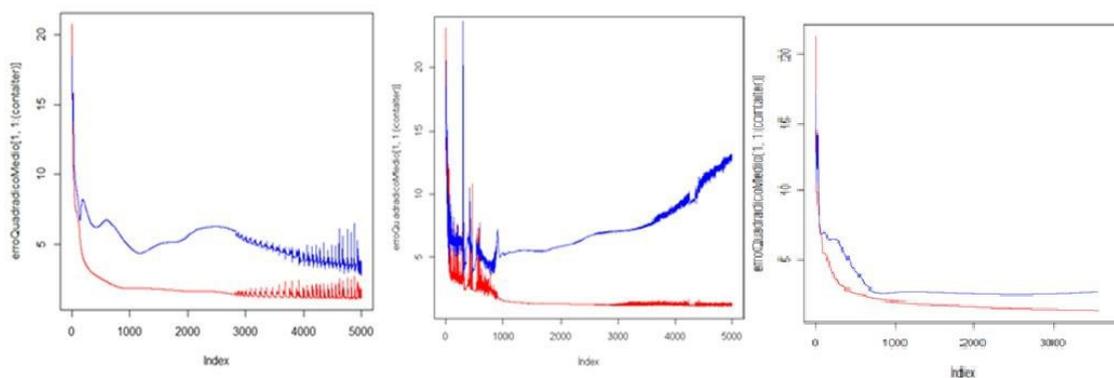


Figura 3. Curvas de MSE para *Backpropagation*, delta-bar-delta e algoritmo LM

As curvas do MSE com menor erro de treinamento e menor erro de validação cruzada são apresentadas, respectivamente em vermelho e azul na Figura 3, para o algoritmo *Backpropagation*, a regra delta-bar-delta e algoritmo LM, nesta ordem. O comportamento dos gráficos da Figura 3 foi predominante nos resultados deste experimento. A curva do erro de treinamento obedece, na maioria das vezes, ao comportamento enunciado pela literatura [Haykin 1999, Fausset 1994, Kovacs 1996].

Como pode-se observar na Figura 3, as curvas do erro do algoritmo *Backpropagation* e do algoritmo LM são similares. Isto significa que praticamente não houve ganho ao executar a otimização prevista pelo algoritmo LM. Isto ocorre pois, como citado na seção 3, a atualização realizada na matriz de pesos da rede neural após o passo *Backward* do algoritmo *Backpropagation* pode ser desfeita nos casos em que o erro torna-se maior. A curva do erro de validação cruzada da regra delta-bar-delta cresce bem mais cedo que as demais. Segundo [Haykin 1999], este é um sinal de que o treinamento pode terminar bem antes, em outras palavras, torna-se perigosa a sua continuidade, sob o risco de perda de generalidade da rede. Isto explica o menor número utilizado para o limite de iterações deste algoritmo em relação ao *Backpropagation*.

Para o teste de Kruskal-Wallis, o p-valor médio obtido ao se comparar o conjunto de MSE dos três tipos de conjuntos de dados (neutro, negativo e positivo) foi 0.89. Ou seja, para um nível de significância de 5%, não se pode rejeitar a hipótese nula H_0 levantada na seção 5.2, e **não** se pode afirmar que os resultados obtidos pelos algoritmo LM e a regra delta-bar-delta, apresentaram resultados melhores que o tradicional *Backpropagation*.

7. Conclusões e trabalhos futuros

Este trabalho apresentou os resultados obtidos ao se treinar diferentes redes neurais utilizando componentes independentes estimadas pelo algoritmo FastICA. O teste de normalidade KS apresentou fortes evidências de que os conjuntos de dados coletados não são gaussianos, condição necessária para execução dos algoritmos em ICA.

O desempenho de todos os algoritmos para separar as componentes foi similar, tanto em número de iterações quanto no MSE calculado nos passos posteriores. Observou-se que o erro quadrático médio das redes neurais utilizadas não difere sob o teste de Kruskal-Wallis. Isto significa que não se pode afirmar em melhor desempenho de um ou outro algoritmo. O fato é que, em todas as configurações para as redes neurais, os algoritmos só pararam ao atingir o número limite de iterações.

Um ponto importante a ser observado é que as componentes estimadas, mesmo separadas, ainda contém sinais que não são de interesse. Então, outros sinais podem estar elevando o valor do MSE obtido, dificultando o aprendizado das redes neurais. Em trabalhos futuros, será pesquisado como reconhecer e extrair componentes de interesse baseando-se na densidade espectral das componentes, e na utilização de filtros. Serão utilizadas novas bibliotecas que permitam outras alternativas de algoritmos para estimar as componentes e treinar as redes neurais.

Referências

- Basili, V., Gianluigi, C., Rombach, H. D. (1994). The Goal Question Metric Approach.
- Basili, V. R., Selby, R.W., Hutchens, D. H. (1986). Experimentation in software engineering, IEEE Transactions in Software Engineering 12 (7) 733-743.
- Corder, G. W., Foreman, D. I. (2009). Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach.
- Comm, P. (1994). Independent component analysis—A new concept? Signal Processing. Vol. 36, pp. 287-314.

- Fausset, L. (1994). Fundamentals of neural network: architectures, algorithms and applications, Prentice Hall.
- Gopu, H., Neelaveni, R., Porkumara, K. (2009) Analysis of EGG Signals for Digestive System Disorders Using Neural Networks.
- Haykin, S. (1999) Neural Networks - A Comprehensive Foundation, Second Edition, Pearson Prentice Hall.
- Hubka, P., Rosik, V., Zidnak, J., Tysler, M., Hulin, I. (2005) Independent Component Analysis of Electrogastrographic Signals. Measurement Science Review, Volume 5.
- Hyvärinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans. Neural Networks, Vol10, pp.626-634.
- Hyvärinen, A., Karhunen, J., Oja, E. (2001). Independent Component Analysis, Wiley Interscience Publication.
- Hyvärinen, A. (1997). A Fast Fixed-Point Algorithm for Independent Component Analysis, Neural Computation.
- Komorowski, D., Pietraszek, S. (2009); Preprocessing for Spectral Analysis of Electrogastrogram. WC 2009, IFMBE Proceedings 25/IV.
- Kovacs, Z. I. (1996). Redes neurais artificiais: fundamentos e aplicações, um texto básico, Editora Livraria da Física.
- Shorack, G. R., Wellner, J. A. (1986). Empirical Processes with Applications to Statistics. Wiley. p. 239.
- Todd, M.J. (1976). The computation of fixed points and applications. Springer, Berlin.
- Wang, Z., Chen, Z. (1997) Blind EGG Separation Using ICA Neural Networks. 19th International Conference - IEEE/EMBS.
- Wu, X., Lu, J., Chen, K., Long, Z., Wang, X., Shu, H., Li, K., Liu, Y., Yao, L. (2009) Multiple neural networks supporting a semantic task.
- Yu, S., Chou, K. (2006) Combining Independent Component Analysis and Backpropagation Neural Network for ECG Beat Classification. Proceedings of the 28th IEEE EMBS Annual International Conference.