EFICIÊNCIA DE DIFERENTES REDES NEURAIS PARA DISCRIMINAR DADOS DE BIOMECÂNICA

[1]Thiago Santana Lemes, [1]Adriano Péricles Rodrigues, [1]Marcus Fraga Vieira

[1]Laboratório de Bioengenharia e Biomecânica/Universidade Federal de Goiás, Goiânia, Brasil

Resumo: Este estudo analisou a eficiência de diferentes algoritmos de máquina de vetor de suporte (SVM) para discriminar dados de diferentes sujeitos. Utilizou-se dados previamente coletados de idosos e jovens com 3 coletas por sujeito, em um estudo de controle postural na plataforma de força. Os dados foram analisados a partir da densidade espectral de potência (PSD) do centro de pressão sobre a qual foi aplicada a análise de componentes principais (PCA) para reduzir a dimensionalidade dos dados. A SVM recebeu a PCA com 90% de variância da PSD original e utilizando diferentes núcleos de produto interno calculou a eficiência de cada um para diferenciar grupos com características distintas. A SVM que obteve o melhor desempenho foi a de núcleo Polinomial, com uma eficiência de 90% aproximadamente, no entanto, o resultado é dependente dos dados a serem classificados, e se faz necessário então uma ferramenta que possa utilizar diferentes núcleos.

Abstract: This study analyze the efficiency of different algorithms of support vector machine (SVM) to discriminate data from different subjects. It was used data previously collected from elderly and young people with 3 collections by subject, in a postural control study on a force plate. Data were analyzed from the power spectral density (PSD) of the center of pressure on which was applied principal component analysis (PCA) to reduce the dimensionality of the data. The SVM received the PCA with 90% of the variance of the original PSD and using different inner product kernels was calculated the efficiency of each one to differentiate between groups with different characteristics. The SVM that have the best performances was the Polynomial with an efficiency of 90% approximately, however, the result depends on data to be classified and it is necessary then a tool that can use different cores.

Descritores: Redes Neurais (Computação); Reconhecimento automatizado de padrão; Processamento de Sinais Assistido por Computador.

Introdução

Algoritmos de classificação são muito utilizados atualmente no intuito de automatizar processos e melhorar análises de dados. As máquinas de vetor de suporte (SVM) são algoritmos que mapeiam vetores de entrada em espaços de características multidimensionais os quais servirão para construir hiperplanos ótimos de separação. SVM's são algoritmos supervisionados que necessitam de amostras de treinamento, usando os resultados desejados, para que o padrão dos dados seja aprendido. Uma amostra de treinamento segue a forma , onde x_i é o padrão de entrada, d_i o resultado desejado e i representa o i-ésimo exemplo de treinamento[1]. Após o treinamento a SVM irá classificar os padrões apresentados em diferentes categorias. Os algoritmos de SVM se diferenciam por seus núcleos de produto interno como, por exemplo, Máquina de Aprendizagem Polinomial, Quadrática, Rede de Função de Base Radial (RBF) e Perceptron de Múltiplas Camadas (MLP) os quais definem diferentes superfícies de decisão não lineares[2].

Na máquina de aprendizagem polinomial, o usuário deve especificar o grau do polinômio a ser utilizado. Se o polinômio tiver grau 2, será então um núcleo quadrático. RBF's definem centros em

regiões circulares baseados na quantidade de vetores de suporte e o usuário define o raio da região circular. MLP's definem vetores de pesos também pelo número de vetores de suporte.

O uso de redes neurais artificiais (RNA) para diferentes propósitos vem sendo amplamente utilizado em biomecânica. Trabalhos como[3,4,5] se utilizam das RNA em diagnósticos, problemas de separação, predição e classificação de dados. Assim, nesse trabalho nosso objetivo foi analisar a eficiência dos diferentes tipos de algoritmos SVM na classificação de dados biomecânicos.

Tais algoritmos são vantajosos por permitirem automatizar e agilizar processos, auxiliando na tomada de decisões para médicos, fisioterapeutas e demais profissionais da área da saúde. Essas tomadas de decisões são importantes em diagnósticos e pesquisas, que, entretanto, demandam muito tempo e carecem de ferramentas automatizadas.

Métodos

Coleta de Dados

Nesse trabalho foram utilizados dados previamente existentes no laboratório, relativos à oscilação do centro de pressão (COP) de 17 idosos e 18 jovens. Foram coletados dados com uma plataforma de força AMTI na posição horizontal e posicionada em aclive e declive com inclinação de 14º. Os dados do COP foram capturados nas situações de olhos abertos e fechados, declive, aclive e horizontal entre jovens e idosos. Os participantes realizaram 3 tentativas em cada situação.

Processamento

Utilizando o software MATLAB, os dados foram processados e, após testes preliminares, foi constatado que a maior diferença entre as situações está no domínio da frequência. Produziu-se, então, uma grande quantidade de dados a partir da densidade espectral de potência (PSD) das oscilações anteroposteriores do COP. Devido à grande quantidade de dados, realizou-se uma análise de componentes principais (PCA) na PSD entre 0.05 e 2.0 Hz para reduzir a dimensionalidade dos dados. Esse intervalo de frequências foi adotado pois 90% da potência total da oscilação do COP é encontrada abaixo de 2.0 Hz[6, 7].

Os dados das oscilações anteroposteriores do COP foram armazenados em uma matriz, e cada linha correspondendo a uma tentativa de um sujeito. A PCA foi aplicada à essa matriz e as componentes principais são as características da forma de onda baseadas na variabilidade da forma de onda original, sendo ortogonais umas às outras. As primeiras componentes principais correspondem às maiores fontes de variabilidade[8].

A PCA foi utilizada pois a grande quantidade de dados dificulta a análise, elevando o tempo de processamento. A PCA reduz a dimensionalidade dos dados, separando os auto-vetores da matriz de dados que contém a maior parte da variância dos dados originais. As vinte primeiras PCA's, foram escolhidas por *scree test*, contendo 90% da variância total dos dados. Na Figura 1 são mostradas as duas primeiras componentes principais como exemplo. Em seguida, as PCA's selecionadas foram inseridas nos diferentes algoritmos de SVM contidos na toolbox do MATLAB.

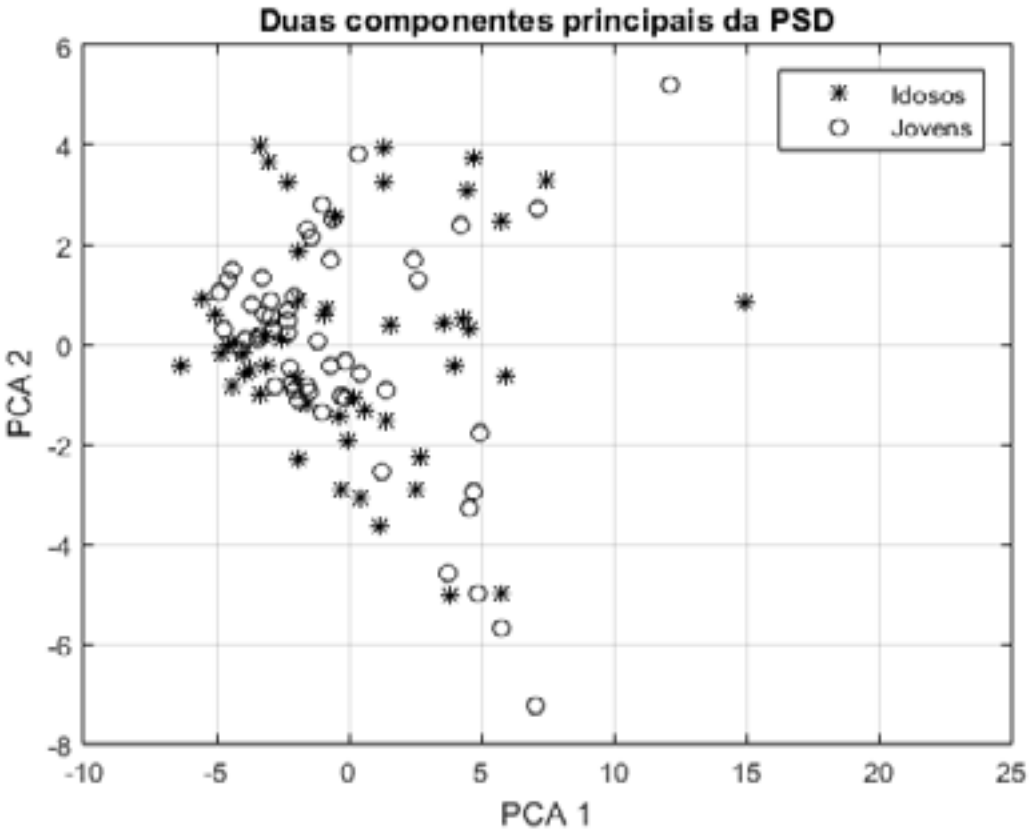

Figura 1: Exemplo de duas componentes principais da PSD na situação aclive de olhos abertos

Algoritmos da SVM

Existem vários tipos de algoritmos de SVM, no entanto o problema básico envolvendo padrões linearmente separáveis requer um hiperplano simples como ilustrado na Figura 2:

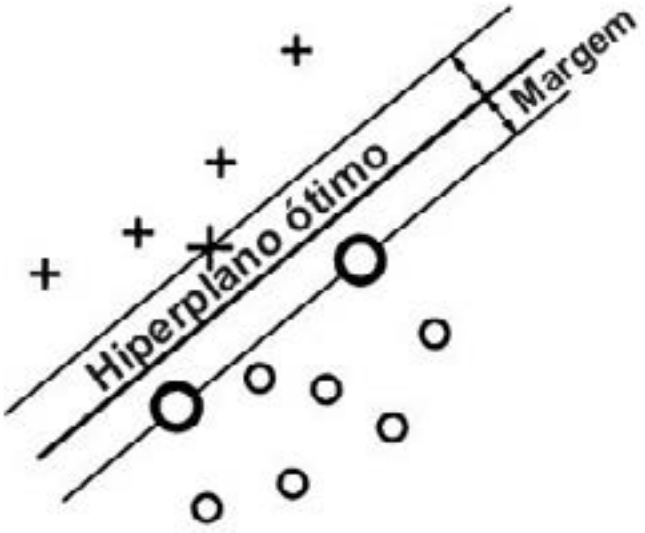

Figura 2: Hiperplano de separação ótimo que separa dados com a máxima margem.[2]

As seguintes expressões definem o hiperplano:

$$w^T x_i + b \geq 0 \quad (1)$$

$$w^T x_i + b < 0 \quad (2)$$

onde *w* representa um vetor de pesos ajustáveis, x_i representa os dados de entrada para o i-ésimo exemplo e *b* é um bias. Os vetores de suporte ("o" e "x" sobre os segmentos de reta na Figura 2) são justamente os valores particulares de *x* que estão no limite do conjunto, ou seja, mais próximos da superfície de decisão, e separam os dados com a máxima margem possível,[1] tomada como a distância entre o hiperplano e os dados mais próximos (Figura 2). Os problemas não linearmente separáveis, que são a maioria, exigem uma abordagem mais complexa, e, portanto, justifica-se a introdução dos núcleos de produto interno, que são funções capazes de obter hiperplanos de separação mais sofisticados.

Inicialmente foi utilizado o algoritmo de SVM com o núcleo de produto interno Polinomial, para a separação entre os grupos jovens x idosos para todos os casos: olhos abertos, fechados, aclive, declive e horizontal. Logo em seguida foi realizada a separação entre as condições olhos abertos x olhos fechados para jovens, idosos, aclive, declive e horizontal. O procedimento se repetiu alterando núcleo de produto interno da SVM para RBF, Quadrática, Linear e MLP.

O núcleo RBF pode ser definido pela seguinte expressão:

$$K(x_i, x_j) = e^{\left(\frac{|x_i - x_j|^2}{2\sigma^2}\right)} \quad (3)$$

na qual *K* é o núcleo do produto interno entre o vetor de entrada x_i e vetor de padrão de entrada x_j e σ representa a largura da região gaussiana e tem valor padrão igual a 1 na toolbox.

O núcleo Linear pode ser definido pela expressão:

$$K(x_i, x_j) = (x_i . x_j) \quad (4)$$

na qual *K* representa o núcleo do produto interno entre o vetor de entrada x_i e o vetor de padrão de entrada x_j.

O núcleo MLP pode ser definido pela seguinte expressão com uso de tangente hiperbólica (*tanh*):

$$K(x_i, x_j) = tanh(\beta_0(x_i . x_j) + \beta_1) \quad (5)$$

K é o núcleo de produto interno, x_i é o vetor de entrada e x_j é o vetor de padrão de entrada. Por padrão na toolbox os pesos têm valor 1 e bias igual a -1. β_0 e β_1 são parâmetros que devem ser respectivamente maior do que zero e menor do que zero. Aqui são assumidos por padrão [+1 -1].

O núcleo Polinomial pode ser definido pela seguinte expressão:

$$K(x_i, x_j) = \left((x_i . x_j) + 1\right)^p \quad (6)$$

na qual *K* representa o núcleo do produto interno, x_i é o vetor de entrada e x_j é o vetor de padrão de entrada, *p* , o grau do polinômio, é um valor inteiro, e por padrão igual a 3 na toolbox.

Para o núcleo Quadrático o valor de *p* é definido como 2 na equação 6.

Em cada caso o objetivo é traçar o hiperplano ótimo que melhor separar os conjuntos, aqui definidos em +1 e -1.

Resultados e Discussão

A Tabela 1 mostra os dados de eficiência de cada núcleo de produto interno. Observa-se que as SVM's com núcleo Polinomial obtiveram a maior eficiência entre todos os testes, atingindo 89.66% de acertos na situação horizontal com olhos abertos, discriminando idosos de jovens. O núcleo MLP, por outro lado, obteve a pior eficiência entre todos os testes, com apenas 46,15% de acertos na situação declive com idosos, separando as condições olhos abertos de olhos fechados.

Tabela 1: Eficiência das diferentes funções núcleo na SVM

	LINEAR	QUAD.	POLI.	MLP	RBF
Idosos vs Jovens					
Declive de olhos abertos	82,76%	65,52%	62,07%	86,21%	72,41%
Declive de olhos fechados	72,41%	65,52%	55,17%	75,86%	51,72%
Horizontal de olhos abertos	86,21%	72,41%	89,66%	72,41%	68,97%
Horizontal de olhos fechados	86,21%	62,07%	75,86%	65,52%	58,62%
Aclive de olhos abertos	82,76%	75,86%	72,41%	79,31%	86,21%
Aclive de olhos fechados	86,21%	75,86%	72,41%	86,21%	72,41%
Média	84,49%	68,97%	72,41%	77,59%	70,69%
Olhos abertos vs Olhos fechados					
Idosos no declive	61,54%	50,00%	50,00%	46,15%	61,54%
Idosos na horizontal	65,38%	50,00%	53,85%	57,69%	61,54%
Idosos no aclive	57,69%	57,69%	53,85%	53,85%	53,85%
Jovens no declive	53,13%	62,50%	62,50%	56,25%	59,38%
Jovens na horizontal	53,13%	68,75%	50,00%	50,00%	53,13%
Jovens no aclive	75,00%	71,88%	68,75%	84,38%	71,88%
Média	59,62%	60,10%	53,85%	55,05%	60,46%

Todavia, analisando a média das eficiências de cada rede em separar as diferentes características, idosos x jovens e olhos abertos x olhos fechados, os resultados obtidos apontam que o núcleo Linear foi melhor para diferenciar idosos de jovens. Entretanto, para diferenciar olhos abertos de olhos fechados o núcleo RBF apresentou uma média melhor, com 60,46% de acerto, embora todos os núcleos tenham obtido uma média próxima ou pouco abaixo de 60%.

A Figura 3 mostra a PSD média anteroposterior de jovens vs idosos na condição horizontal de olhos abertos (em cima), e olhos abertos vs fechados de idosos no declive (embaixo), que foram as condições nas quais foram obtidas a melhor e a pior eficiência das SVM's, respectivamente.

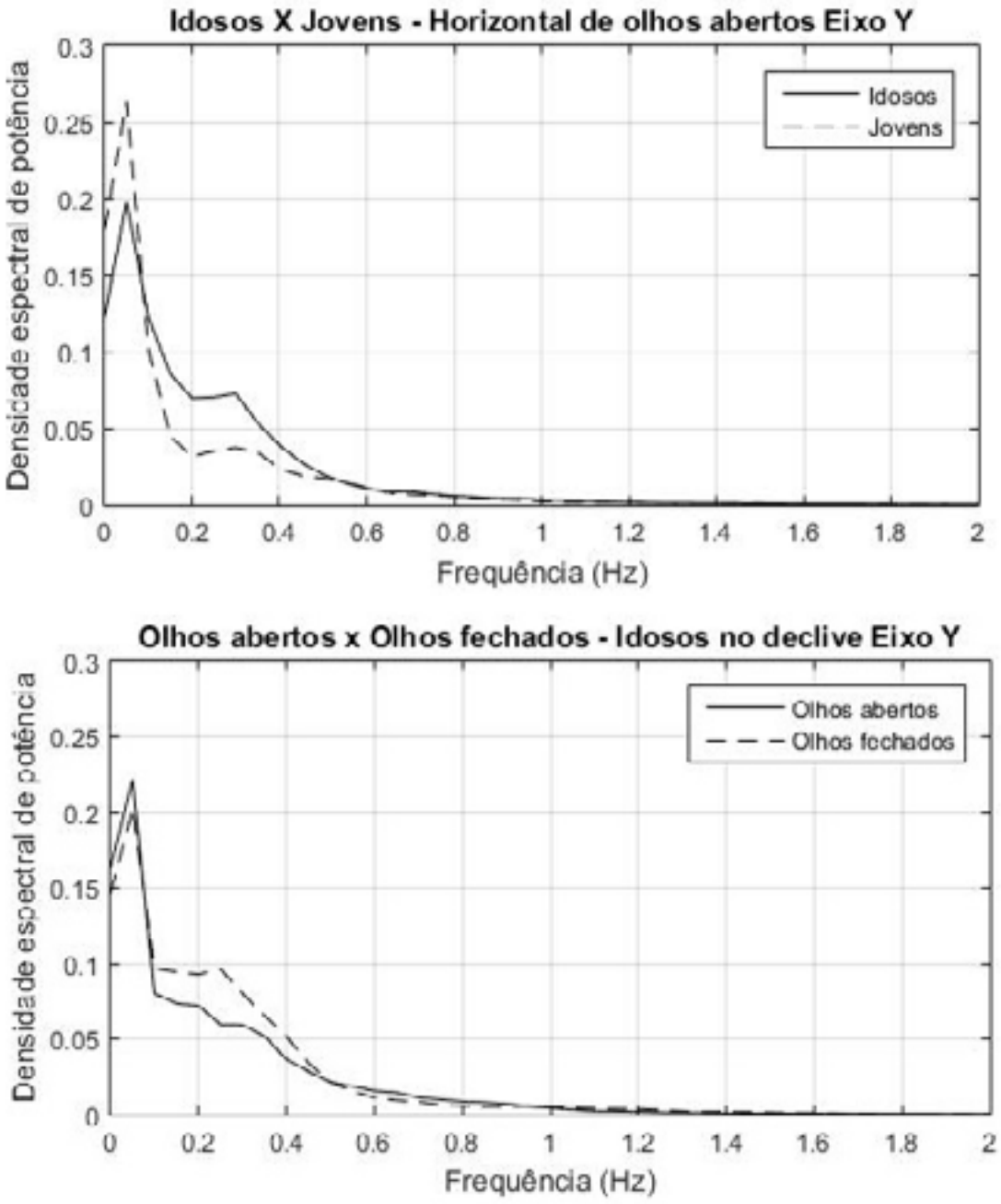

Figura 3**:** Densidade espectral de potência do melhor (em cima) e pior (embaixo) resultado

Na condição horizontal com olhos abertos, as curvas médias são visivelmente separáveis o que justifica o bom resultado obtido com a SVM de núcleo Polinomial. Na condição declive com idosos de olhos abertos vs fechados, por outro lado, a separação entre as curvas médias não é tão evidente, o que justifica o resultado insatisfatório obtido com o núcleo MLP.

Como mostrado na Tabela 1, para os dados analisados, o núcleo de produto interno que obteve o melhor desempenho foi o Polinomial, atingindo valores acima de 89% em determinadas condições. Desempenho semelhante pode ser visto no trabalho de Wu[9], que apesar de utilizar um método diferente de avaliação da SVM obteve 90% de eficiência com esse mesmo núcleo classificando dados biomecânicos da marcha.Barton e Lees[10] obtiveram um desempenho de 83,3% utilizando uma rede neural do tipo MLP para diferenciar padrões da marcha de 8 indivíduos caminhando em esteira, um resultado contrastante ao obtido no presente estudo.

Por outro lado, Muniz e seus colegas[3] obtiveram o melhor desempenho na classificação de seus dados com a SVM utilizando um núcleo de produto interno Linear, o que se assemelha com os resultados obtidos pelas médias no presente estudo, levantando a hipótese de que o desempenho da SVM depende, além de seus núcleos de produto interno, dos dados de entrada.

A Figura 4 mostra a separação dos dados de teste em um diagrama das duas primeiras componentes principais selecionadas (PC1 x PC2). Observa-se que uma separação linear dos dados não seria adequada. Porém, levando em consideração as 20 componentes principais utilizadas, a separação linear se torna mais eficiente, diferente do trabalho de Barton e Lees[10], o que sugere que os dados influenciam na escolha do núcleo de produto interno da SVM, bem como o número de dados fornecidos para a SVM.

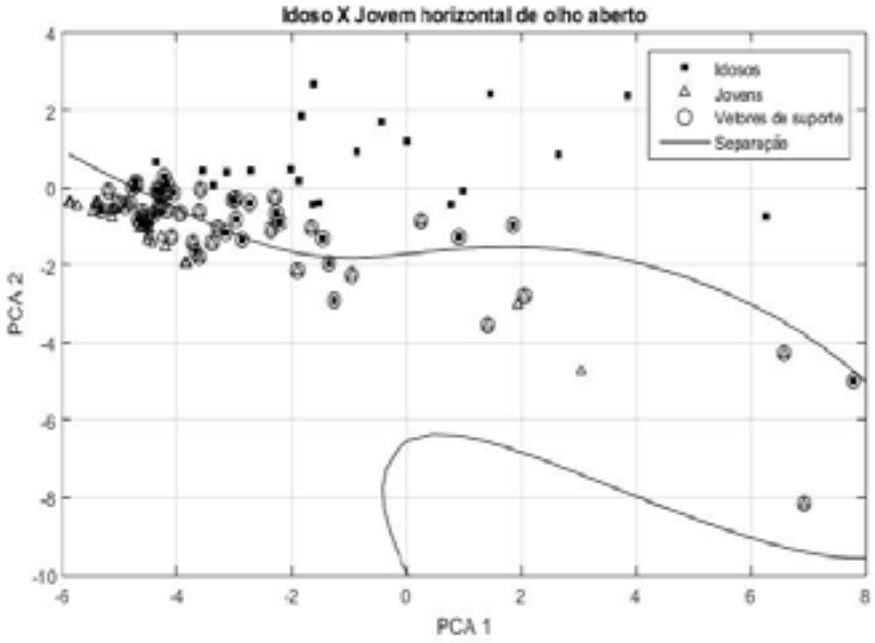

Figura 4: Eficiência na separação dos dados da SVM com função núcleo Polinomial

De acordo com Rauber e Berns[11], os resultados numéricos não justificam quaisquer conclusões teóricas sobre o núcleo de produto interno, mas apontam informações interessantes sobre os dados.

Assim, cada situação deve ser cuidadosamente analisada a fim de testar a eficiência de cada núcleo.

Conclusão

Baseado nos dados coletados e nas análises realizadas, o que se pode afirmar é que a escolha do núcleo de produto interno depende diretamente de quais são os dados a serem analisados. Desta forma, sugere-se que um sistema automatizado para reconhecimento de padrões em biomecânica contenha várias opções de núcleo para RNA.

Agradecimentos

Agradecemos à CAPES pela bolsa de estudos à Thiago Santana Lemes e Adriano Péricles e ao CNPQ pelo apoio financeiro ao projeto.

Referências

[1] HAYKIN SIMON. Rede Neurais Princípios e Prática. 2ª Edição. Bookman: Porto Alegre; 2001.
[2] VAPNIK VLADIMIR N. The Nature of Statistical Learning Theory. 2nd ed. Springer: New York; 2000.
[3] MUNIZ A.M.S, LIU H, LYONS K.E, PAHWA R, LIU W, NOBRE F.F, et al. Comparison among probabilistic neural network, support vector machine and logistic regression for evaluating the effect of subthalamic stimulation in Parkinson disease on ground reaction force during gait. Journal of Biomechanics. 2010; 43: 720–726.
[4] KACZMARCZYK K, WIT A, KRAWCZYK M, ZABORSKI J, PILSUDSKI J. Artificial Neural Networks (ANN) Applied for Gait Classification and Physiotherapy Monitoring in Post Stroke Patients, Artificial Neural Networks - Methodological Advances and Biomedical Applications, Prof. Kenji Suzuki (Ed.),
ISBN: 978-953-307-243-2, InTech, Available from: http://www.intechopen.com/books/artificial-neural-networksmethodological-advances-and-biomedical-applications/artificial-neural-networks-ann-applied-for-gaitclassification-and-physiotherapy-monitoring-in-post.
[5] DE VRIES W.H.K, VEEGER H.E.J, BATEN C.M.T, VAN DER HELM F.C.T. Determining a Long Term Ambulatory Load Profile of the Shoulder Joint: Neural Networks Predicting Input for a Musculoskeletal Model. Human Movement Science. 2012; 31:419-428.
[6] MEZZARANE R.A, KOHN A.F. Control of upright stance over inclined surfaces. Exp Brain Res. 2007; 180:377-388.
[7] HAYES K.C. Biomechanics of postural control. Exerc Sport Sci Rev. 1982; 10:363-391.
[8] MUNIZ A.M.S, NADAL J, LYONS K.E, PAHWA R, LIU W. Long-term evaluation of gait initiation in six Parkinson's disease patients with bilateral subthalamic stimulation. Gait & Posture. 2012; 35:452-457.
[9] WU JIANNING, WU BIN. The Novel Quantitative Technique for Assessment of Gait Symmetry Using Advanced Statistical Learning Algorithm. BioMed Research International. 2015.
[10] BARTON J.G, LEES A. An application of neural networks for distinguishing gait patterns on the basis of hip-knee joint angle diagrams. Gait & Posture. 1997; 28-33.
[11] RAUBER THOMAS W, BERNS KARSTEN. Kernel Multilayer Perceptron. Proceedings of the 2011 24th SIBGRAPI Conference on Graphics, Patterns and Images; 2011 Aug 28-31; Maceió, Brasil. 2011. p. 337–343.

Contato

Thiago Santana Lemes
Mestrando em Engenharia Elétrica e de Computação, Laboratório de Bioengenharia e Biomecânica-Universidade Federal de Goiás – UFG. Campus Samambaia, Avenida Esperança s/n, CEP: 74.690-900. Goiânia – Goiás - Brasil
(62)98422-9667
thiago2608santana@gmail.com