Anais do V Congresso Brasileiro de Eletromiografia e Cinesiologia e X Simpósio de Engenharia Biomédica - ISBN: 978-85-5722-065-2 - DOI: 10.29327/cobecseb.78871

ÁRVORES DE DECISÃO COMO FERRAMENTA DE AUXÍLIO NA ANÁLISE BIOMECÂNICA

A. Carafini[1], F. B. Rodrigues[1], G. A. G. De Villa[1], T. S. Lemes[1], A. O. Andrade[2] e M. F. Vieira[1]

[1] Laboratório de Bioengenharia e Biomecânica, Universidade Federal de Goiás, Goiânia, Brasil
[2]Núcleo de Inovação e Avaliação Tecnológica em Saúde (NIATS), Faculdade de Engenharia Elétrica, Universidade Federal de Uberlândia, Uberlândia, Brasil
e-mail: carafini.adriano@gmail.com

Resumo: Um exemplo de aplicação de árvores de decisão como ferramenta de auxílio na análise biomecânica é apresentado. O estudo foi conduzido com séries temporais de componentes verticais da força de reação do solo coletadas de 6 amputados unilaterais transfemorais e 6 indivíduos sem amputação, durante a marcha, em situações de declive e horizontal. As regras de decisão foram tomadas para obtenção das características temporais de marcha dos transfemorais que mais se diferem da marcha normal. Nas duas situações, o tempo do primeiro pico de força do lado protético foi o principal discriminante, obtido com taxas médias de acerto de 91.7% e 96.9% nas situações horizontal e declive, respectivamente.
Palavras-chave: árvores de decisão, marcha, transfemoral.

Abstract: *An example of decision trees application as an auxiliary tool for biomechanical analysis is presented. The study was conducted over temporal series of vertical ground reaction force collected from 6 transfemoral amputee and other 6 control individuals, in plane and inclined situations. The decision rules were taken as basis for extraction of the temporal characteristics of the transfemoral gait that most differed from a normal gait. In both situations, the first force peak time on the prosthetic side was the main discriminant, obtained with mean accuracies of 91.7% and 96.9% in horizontal and inclined situations, respectively.*
Keywords: *decision tree, gait, transfemoral.*

Introdução

A análise de marcha de amputados unilaterais transfemorais tem sido estudada por vários autores [1–5]. Dundass et al. avaliou os efeitos da deterioração do controlador de joelho hidráulico [1], e Van der Linden os efeitos de vários tipos de pés prostéticos [5] na marcha de amputados unilaterais transfemorais. Jaegers et al. analisou parâmetros cinemáticos, bem como padrões de movimento do tronco, do quadril e joelho, encontrando evidências de que amputados apresentam padrões assimétricos na marcha [2]. Estes mesmos autores observaram, ainda, que transfemorais apoiam por mais tempo sobre a perna intacta.

Tipicamente, dados biomecânicos apresentam alta variabilidade inter e intra-indivíduos, interação não linear entre alguns parâmetros e multidimensionalidade, o que causa dificuldade na análise de alguns casos particulares através de métodos convencionais [6,7]. Além disso, o sucesso da análise é fortemente limitado pela habilidade e experiência do investigador, que deve ser capaz de lidar com grandes conjuntos de dados e fazer inferências baseadas em seu conhecimento prévio acerca do objeto de análise [6]. Uma forma de superar essa limitação é a aplicação de técnicas de aprendizagem de máquina, tais como redes neurais, árvores de decisão, máquina de vetor de suporte, etc.

A exploração de dados biomecânicos através de árvores de decisão já tem sido realizada. Derlatka e Ihnatouski analisaram e classificaram dados de pressão plantar de indivíduos que sofrem de pé valgo e paralisia cerebral [6], e Heller et al. reconstruíram os padrões de ativação muscular a partir de dados cinemáticos [8]. O uso de árvores de decisão tem como vantagem a facilidade de extração de informação acerca dos dados devido ao formato simples das regras de decisão resultantes do seu processo de aprendizado, que, ao apontarem os principais discriminantes, reduzem a complexidade de interpretação dos dados, mitigando possíveis limitações impostas pela experiência *a priori* do investigador.

Neste trabalho, a técnica de árvores de decisão foi empregada na análise de marcha de amputados unilaterais transfemorais, nas situações horizontal (sem inclinação) e declive (8%), a partir de parâmetros temporais do perfil da componente vertical da força de reação do solo (VFRS). O objetivo é apresentar uma aplicação das árvores de decisão e a interpretação de seus resultados do ponto de vista biomecânico.

Materiais e métodos

Participaram deste estudo um total de 12 sujeitos (Tabela 1), sendo 6 participantes do grupo controle (CT) e 6 amputados unilaterais transfemorais (TF). Todos os TF possuem pelo menos um ano de experiência com a prótese, deambulam sem auxílio de qualquer equipamento e suas amputações tem origem traumática.

As séries temporais da VFRS força plantar foram coletadas em 8 repetições de 30 segundos, enquanto cada indivíduo caminhava em uma esteira ergométrica com

velocidade auto-selecionada (VAS), previamente identificada.

Tabela 1. Caracterização dos grupos amostrais.

	Controle	**Transfemoral**
N / sexo	6 ♂	6 ♂
Idade (anos)	32,20 ± 10,17	32,38 ± 8,65
IMC (kg/cm^2)	24,69 ± 3,10	23,60 ± 2,63
Tempo amputação (anos)	NA	11,11 ± 6,90
Tempo protetização (anos)	NA	9,61 ± 5,19
Comprimento do coto (cm)	NA	36,08 ± 7,83
VAS (km/h)	4,94 ± 0,74	3,27 ± 0,69

Palmilhas de sensores de pressão F-Scan (Tekscan, Inc.307 West Firts Street, South Boston, MA0, USA), devidamente ajustadas dentro dos calçados (calçado de uso cotidiano) de cada participante, foram utilizadas na aquisição dos dados a uma taxa de amostragem de 100 Hz.

Os sinais de VFRS foram obtidos a partir das sequências de mapas de pressão utilizando o *software* F-Scan e foram filtradas por um filtro passa-baixa *Butterworth* de fase zero, com ordem efetiva igual a 8 e frequência de corte em 6 Hz [1]. Rotinas no Matlab (2015a) foram desenvolvidas para a extração de 5 parâmetros temporais (dos lados direito e esquerdo), sendo eles o tempo de duplo apoio (TDA), tempo de apoio (TA), tempo de primeiro pico de força (TPP), tempo de vale (TV), tempo de segundo pico (TSP). A Figura 1 ilustra os pontos tomados como referência para extração dos parâmetros.

Os tempos de duplo apoio foram separados entre TDA do lado protético e do lado intacto. O TDA do lado protético, por exemplo, foi tomado como sendo o tempo entre o início do apoio do lado protético e fim do apoio do lado intacto.

As diferenças relativas, calculadas a partir da Equação 1, mitigam a influência das diferenças de velocidade de marcha, e evidenciam as assimetrias entre os dois lados. Logo, as diferenças relativas entre os parâmetros da perna protética e intacta foram utilizadas como entradas do algoritmo de árvore de decisão do módulo de aprendizagem de máquina *scikit-learn* (Python). Esse módulo foi escolhido pela riqueza de informações e a forma concisa em que são apresentadas em sua saída (Figura 2).

$$P = \frac{\mathrm{P_{protético}} - P_{intacto}}{P_{protético}} \quad (1)$$

Onde *P* representa um parâmetro qualquer.

Para cada grupo e situação (horizontal e declive), as amostras foram divididas em três segmentos e, para cada classificação, dois terços foram utilizados como amostras de treino e o restante para validação [9]. Então, apenas uma árvore de decisão para cada situação foi obtida consumindo todas as amostras, de modo a sintetizar os principais discriminantes. Por fim, o mesmo processo de classificação foi realizado excluindo do grupo de variáveis os principais discriminantes obtidos na primeira aplicação do método de classificação.

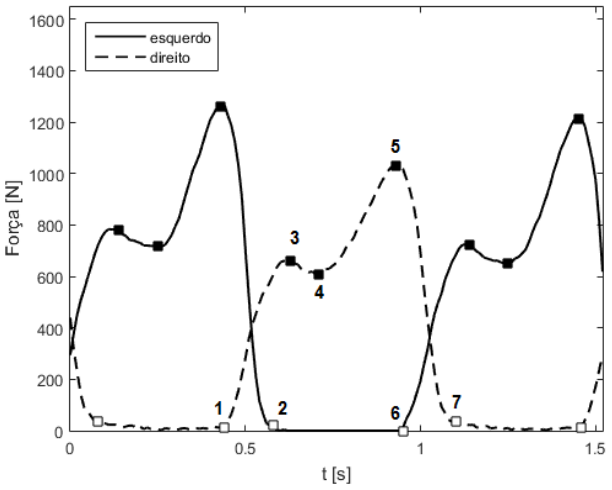


Figura 1: Força vertical (pé direito e esquerdo) de um dos indivíduos do grupo controle, na situação horizontal. Os parâmetros TPP, TV e TSP são dados, por exemplo, pelos intervalos 1-3, 1-4 e 1-5, respectivamente. O parâmetro TA é dado pelo intervalo 1-7. Os parâmetros TDA do lado direito e esquerdo são dados, respectivamente, pelos intervalos 1-2 e 6-7.

A profundidade de cada árvore de decisão foi determinada como aquela que maximiza as probabilidades de classificação correta de ambas as classes e minimiza o *overfitting* sobre as amostras de treino, uma vez que baixas probabilidades de acerto e/ou *overfitting* exacerbado sobre as amostras de treino podem implicar em resultados não generalizáveis.

Resultados

A Figura 2 ilustra as árvores de decisão obtidas para as duas situações (horizontal e declive), ao se utilizar todas as amostras no processo de aprendizagem. As taxas médias de acerto entre as 3 combinações de amostras foram de 91.7% e 96.9%, nas situações horizontal e declive, respectivamente.

Os resultados para o processo de análise excluindo a diferença relativa de TPP, são apresentados pelas decisões em formato de regras '*if-then*'. As taxas médias de acerto entre as 3 combinações de amostras foram de 79.2% e 83.3%, nas situações horizontal e declive, respectivamente.

Horizontal:

- **If** (TA ≤ -0.0067) **and** (TV ≥ 0.1196) **then** Prob.(TF) = 1
- **If** (TA ≥ -0.0067) **then** Prob.(CT) = 0.94

Declive:

- **If** (TA ≤ -0.0411) **then** Prob.(TF) = 0.97

- **If** (TA ≥ -0.0411) **and** (TDA ≤ 0.086) **then** Prob.(CT) = 0.95

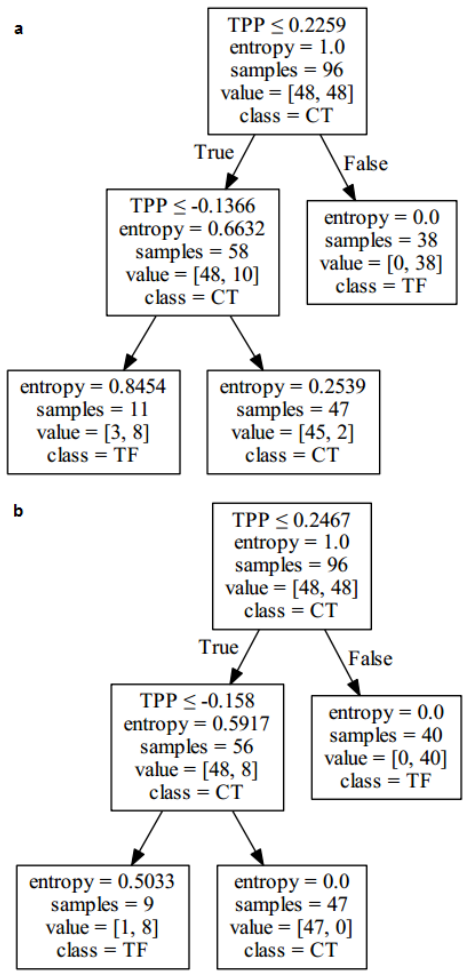


Figura 2: Árvores de decisão final para situação (**a**) horizontal e (**b**) declive, utilizando todas as variáveis na aprendizagem

Discussão

Os resultados ilustrados pela Figura 2 apontam a diferença relativa de TPP como principal e segundo melhor discriminante em todos os casos. Esse atributo utilizado como primeiro nó sugere que os indivíduos com amputação transfemoral demoram um pouco mais a descarregar toda a carga sobre o lado protético, quando comparado com o lado intacto.

O segundo procedimento de aplicação de árvores de decisão, aponta a diferença relativa de TA como principal discriminante. O segundo melhor discriminante não foi o mesmo entre as duas situações, sendo que na condição horizontal foi a diferença relativa de TV e, no declive, foi a diferença de TDA.

O tempo de apoio inferior no lado protético, apontada pelo sinal negativo do *threshold* de decisão, sugere que o amputado transfemoral apoia por menos tempo no lado protético, corroborando com o que foi observado por Jaeger et al. [2]. Ademais, o *threshold* de TA, na situação declive, quase 10 vezes maior em magnitude, quando comparado com o mesmo *threshold* na situação horizontal, aponta que a diferença de tempo de apoio é acentuada na condição de declive.

O tempo tardio de vale, além de ser uma consequência do tempo tardio de primeiro pico, indica que o intervalo de execução de *push-off* é reduzido, sugerindo uma transição mais abrupta da fase de apoio para fase de balanço, no lado protético. Ademais, o TV tardio também é a manifestação de uma estratégia de compensação pela ausência do joelho. O tempo de duplo apoio maior no lado protético justifica o tempo tardio de primeiro pico, e o TDA inferior no lado intacto também corrobora com a sugestão de transição abrupta entre apoio e balanço, no lado protético.

As taxas médias de acerto indicam o quanto as regras de decisão conseguem generalizar para amostras não vistas no processo de treinamento, e servem como índice de validade dos discriminantes escolhidos pelas árvores de decisão. A escolha de um conjunto de regras que leva a baixas taxas de acerto sobre novas amostras, ou incapacidade de generalizar, indica que os discriminantes escolhidos não são representativos de toda a população amostral, mas apenas um resultado viciado do classificador, que tenta maximizar a taxa de acerto apenas sobre as amostras de treino. As taxas de acerto superiores na situação de declive indicam que a complexidade de discriminação entre os dois grupos foi menor e que, conseguintemente, as diferenças entre eles se acentuaram, sugerindo que essa situação impôs maior dificuldade de execução para ambos os grupos.

A Figura 2.b ilustra um exemplo de *overfitting* das amostras de treino, em que uma segunda divisão utilizando a diferença relativa de TPP é realizada para obter duas folhas, sendo que uma delas possui apenas 9 amostras, das 56 contidas no nó pai, e alta entropia, resultando baixo potencial de classificar corretamente novas amostras.

Por fim, vale ressaltar que o tipo de análise proposto nesse trabalho se difere de técnicas convencionais, pois pode ser aplicada a grandes conjuntos de dados sem a necessidade de experiência acerca do objeto de análise, por parte do investigador. Além disso, permite ao investigador a condução de uma análise mais imparcial e automática dos dados.

Conclusão

As análises realizadas sobre os resultados obtidos das classificações por árvores de decisão exemplificam a simplicidade de interpretação das regras de decisão do ponto de vista biomecânico. Essa ferramenta também permite a análise concomitante de um número arbitrariamente maior de variáveis, com um grau de complexidade de interação também arbitrário, já que as regras de decisão apontariam diretamente os principais discriminantes a serem interpretados pelo investigador.

Agradecimentos

O presente trabalho foi realizado com apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento

de Pessoal de Nível Superior (CAPES), Fundação de Amparo à Pesquisa do Estado de Goiás (FAPEG) e Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG). A. O. Andrade é Bolsista do CNPq, Brasil (305223/2014-3).

Referências

[1] C. Dundass, G. Yao, C. Mechefske, Initial biomechanical analysis and modeling of transfemoral amputee gait, J. Prosthetics Orthot. 15 (2003) 20–26. doi:10.1097/00008526-200301000-00006.

[2] S.M.H.J. Jaegers, J.H. Arendzen, H.J. de Jongh, Prosthetic gait of unilateral transfemoral amputees: A kinematic study, Arch. Phys. Med. Rehabil. 76 (1995) 736–743. doi:10.1016/S0003-9993(95)80528-1.

[3] J.L. Stein, W.C. Flowers, Stance phase control of above-knee prostheses: Knee control versus SACH foot design, J. Biomech. 20 (1987) 19–28. doi:10.1016/0021-9290(87)90263-6.

[4] A. Gitter, J. Czerniecki, M. Meinders, Effect of prosthetic mass on swing phase work during above-knee amputee ambulation., Am. J. Phys. Med. Rehabil. 76 (1997) 114–21. http://www.ncbi.nlm.nih.gov/pubmed/9129517 (accessed July 24, 2017).

[5] M.L. Van Der Linden, S.E. Solomonidis, W.D. Spence, N. Li, J.P. Paul, A methodology for studying the effects of various types of prosthetic feet on the biomechanics of trans-femoral amputee gait, J. Biomech. 32 (1999) 877–889. doi:10.1016/S0021-9290(99)00086-X.

[6] M. Derlatka, M. Ihnatouski, Decision Tree Approach to Rules Extraction for Human Gait Analysis, in: Springer, Berlin, Heidelberg, 2010: pp. 597–604. doi:10.1007/978-3-642-13208-7_74.

[7] T. Chau, A review of analytical techniques for gait data. Part 1: Fuzzy, statistical and fractal methods., Gait Posture. 13 (2001) 49–66. http://www.ncbi.nlm.nih.gov/pubmed/11166554 (accessed July 24, 2017).

[8] B.W. Heller, P.H. Veltink, N.J. Rijkhoff, W.L. Rutten, B.J. Andrews, Reconstructing muscle activation during normal walking: a comparison of symbolic and connectionist machine learning techniques., Biol. Cybern. 69 (1993) 327–35. http://www.ncbi.nlm.nih.gov/pubmed/8218535 (accessed July 24, 2017).

[9] T. Hastie, R. Tibshirani, J. Friedman, Cross-Validation, in: Elem. Stat. Learn., 2nd ed., Springer, 2008: pp. 241–249.