World Academy of Science, Engineering and Technology
International Journal of Nutrition and Food Engineering
Vol:16, No:03, 2022

Advancing the Analysis of Physical Activity Behaviour in Diverse, Rapidly Evolving Populations: Using Unsupervised Machine Learning to Segment and Cluster Accelerometer Data

Authors : Christopher Thornton, Niina Kolehmainen, Kianoush Nazarpour

Abstract : Background: Accelerometers are widely used to measure physical activity behavior, including in children. The traditional method for processing acceleration data uses cut points, relying on calibration studies that relate the quantity of acceleration to energy expenditure. As these relationships do not generalise across diverse populations, they must be parametrised for each subpopulation, including different age groups, which is costly and makes studies across diverse populations difficult. A data-driven approach that allows physical activity intensity states to emerge from the data under study without relying on parameters derived from external populations offers a new perspective on this problem and potentially improved results. We evaluated the data-driven approach in a diverse population with a range of rapidly evolving physical and mental capabilities, namely very young children (9-38 months old), where this new approach may be particularly appropriate. Methods: We applied an unsupervised machine learning approach (a hidden semi-Markov model - HSMM) to segment and cluster the accelerometer data recorded from 275 children with a diverse range of physical and cognitive abilities. The HSMM was configured to identify a maximum of six physical activity intensity states and the output of the model was the time spent by each child in each of the states. For comparison, we also processed the accelerometer data using published cut points with available thresholds for the population. This provided us with time estimates for each child's sedentary (SED), light physical activity (LPA), and moderate-to-vigorous physical activity (MVPA). Data on the children's physical and cognitive abilities were collected using the Paediatric Evaluation of Disability Inventory (PEDI-CAT). Results: The HSMM identified two inactive states (INS, comparable to SED), two lightly active long duration states (LAS, comparable to LPA), and two short-duration high-intensity states (HIS, comparable to MVPA). Overall, the children spent on average 237/392 minutes per day in INS/SED, 211/129 minutes per day in LAS/LPA, and 178/168 minutes in HIS/MVPA. We found that INS overlapped with 53% of SED, LAS overlapped with 37% of LPA and HIS overlapped with 60% of MVPA. We also looked at the correlation between the time spent by a child in either HIS or MVPA and their physical and cognitive abilities. We found that HIS was more strongly correlated with physical mobility (R^2HIS =0.5, R^2MVPA= 0.28), cognitive ability (R^2HIS =0.31, R^2MVPA= 0.15), and age (R^2HIS =0.15, R^2MVPA= 0.09), indicating increased sensitivity to key attributes associated with a child's mobility. Conclusion: An unsupervised machine learning technique can segment and cluster accelerometer data according to the intensity of movement at a given time. It provides a potentially more sensitive, appropriate, and cost-effective approach to analysing physical activity behavior in diverse populations, compared to the current cut points approach. This, in turn, supports research that is more inclusive across diverse populations.

Keywords : physical activity, machine learning, under 5s, disability, accelerometer
Conference Title : ICBNPA 2022 : International Conference on Behavioral Nutrition and Physical Activity
Conference Location : Barcelona, Spain
Conference Dates : March 03-04, 2022