A collection of sport activity files for data analysis and data mining 2016a

Authors:

- Samo Rauter (University of Ljubljana)
- Iztok Fister Jr. (University of Maribor)
- Iztok Fister (University of Maribor)

Contact details:

In case of any questions, please let us know at:

- samo.rauter@gmail.com or
- iztok.fister1@um.si

Donation details:

We invite everyone who is interested to donate data to this dataset to contact us.

Download details:

Dataset is available for download as a torrent file via: http://www.academictorrents.com/ or direct LINK

Motivation

The increased popularity of mobile smartphones and smart watches allow people to train smarter. More and more cyclists use a variety of sports applications, such as Strava, Endomondo, Garmin Connect, etc. Those present an ideal tool for monitoring, tracking and analysing their workouts. Each workout can also be exported as an XML file and be analysed [4]. Sports activities offer immense different tasks for data analysis and data mining. In line with this, some ideas were discussed in paper [2]. In future, these applications may also be intended for planning the further sport training sessions, as discussed in [1, 3]

Therefore, some literature about data research in sport were reviewed by us. We realized lack of research data of sport activities and therefore decided to collect a capacity of training data by different outdoor athletes. The majority of them present cyclists with a large amount of experiences and sport involvement.

Description of this dataset

Dataset consists of seven cyclists, who upload their activities to Strava and Garmin Connect profiles. Typically, these activities can be downloaded as a GPX format, which basically presents an XML format. Following features of each training can be extracted: GPS location, elevation, duration, distance, heart rate and even power.

Ethics statement

Data in this dataset were voluntarily donated by cyclists. These cyclists would like to stay anonymous. Since these data is intended for research purposes, it is prohibited for competitors and rivals of these volunteers to analyse their data for own comparison.

References

[1] Iztok Fister, Samo Rauter, Xin-She Yang, Karin Ljubič, and Iztok Jr. Fister. Planning the sports training sessions with the bat algorithm. *Neurocomputing*, 149:993–1002, 2015.

[2] Iztok Jr. Fister, Dušan Fister, Iztok Fister, and Simon Fong. Data mining in sporting activities created by sports trackers. In *Computational and Business Intelligence (ISCBI), 2013 International Symposium on*, pages 88–91. IEEE, 2013.

[3] Iztok Fister Jr, Karin Ljubič, Ponnuthurai Nagaratnam Suganthan, Matjaž Perc, and Iztok Fister. Computational intelligence in sports: Challenges and opportunities within a new research domain. *Applied Mathematics and Computation*, 262:178–186, 2015.

[4] Samo Rauter, Iztok Fister, and Iztok Fister Jr. How to deal with sports activity datasets for data mining and analysis: Some tips and future challenges. *International Journal of Advanced Pervasive and Ubiquitous Computing (IJAPUC)*, 7(2):27–37, 2015.