Journal of Science & Cycling
Breakthroughs in Cycling & Triathlon Sciences

OPEN ACCESS

Abstract

A Gap in the Education of Future Sport Scientists?

N N[1,*], N N[1] and N N[1]

[1] **University.** {firstname.lastname}@

* Correspondence: email.

Received: date; Accepted: date; Published: date

Keywords: data, technology, education

1. Introduction

Technological advances of the last decades have seen a vast increase in data availability in general but also in particular in the domain of sport science. Given the amount of data produced in cycling and other endurance sports, various models have been developed for (among other aspects) predicting performances based on historical data.

While the complexity of data analysis tasks might not have changed (or even has decreased), the possibilities have drastically increased over the last years. This is again also due to the increasing amount of data collected. With the emergence of low-cost, energy-efficient GPS head units for cycling and the ability to store and share data online, the amount of data produced by athletes has increased drastically. Recent years have seen a big increase in available health-/sports-related data due to (among other factors) the introduction of 24/7 monitoring devices. Not only is managing and making use of these datasets challenging for non-technicians, but also the increase in data availability raised the complexity of some novel models (e.g., "deep learning") thus also elevating possibilities in analysis, which cannot be assumed to be understandable.

While all this possibly equips sport scientists, researchers, coaches, and also athletes with many possibilities (e.g., monitoring, modelling, ...) the question to be addressed is: are current and future sport scientists prepared for this task?

2. "Data Science"?

In recent years "data science" has become a "hot topic" and is perceived as a field of research and study on its own (De Veaux et al., 2017). Simply speaking, data science examines and develops methods for extracting information and knowledge from data (Dhar, 2013).

As pointed out previously, the rising amount of data not only increases the possibilities for insights into performance improvement, but also raises the bar with respect to the required data science competencies and skills required for performing an analysis. Demands on people designing such analytical tools rise even further, when the analysis should be automatically executable as soon as a new or updated data set is available and involve little or no effort by the data provider (i.e., the person capturing data during training). While commercial tools and platforms such as TrainingPeaks, WKO+, Today's Plan, Strava, etc. have addressed this issue and often allow coaches and researchers to gain insights into the performances of their athletes, they are limited to a certain amount of predefined metrics and analyses. They do not offer the possibility to extended and adapt analyses according to individual needs or interests. Open-source software such as Golden Cheetah, on the other hand, can potentially be adapted in order to fit individual needs with respect to different types of available analysis capabilities and their depth. However, the skills required for extending the functionality using built-in programming interfaces or possibly extending the functionality of the software by modifying the source code often exceed the technical skill levels that can be expected from sport scientists.

Furthermore, especially when testing novel sensors or models, existing software often does not fit the requirements for implementing the necessary testing protocols. Consequently, ideas are either not realised due to lack of possibilities or are outsourced to software developers who in turn (often) do not have the required domain knowledge in sport science. As a consequence, there is a high chance of missing features or severe analytical errors in the software due to the missing domain knowledge.

In order to mitigate such problems, (future) sport scientists should develop skills and competencies in data science as part of their professional training. A non-exhaustive list of competencies relevant for a rigorous data analysis could be along the lines of the following topics:

- Understanding sensor Technology
- (Mathematical) Modelling
- Visualisation of complex data (more than just creating a standard Excel chart)
- Data processing using a programming language

This list promotes a broad understanding of "data science". For example, "understanding sensor technology" is usually not covered in definitions of data science. However, a thorough understanding of what can be measured and how this process works is required in order to be able to work with data.

3. A Gap in Education?

While it is evident that sport scientists need at least some degree of education in topics related to "data science", there is a gap in the curricula for sport sciences for them. A previous study examined the curricula of sport science universities in Austria revealing revealed that only one university out of four provided students with introductory courses on these topics (Dobiasch & Oppl, 2020).

This finding might be generalisable as also an examination of the undergraduate programs of the Top 3 universities in the "2020 Global Ranking of Sport Science Schools and Departments" (University of Copenhagen, Norwegian School of Sport Sciences and Deakin University) (*ShanghaiRanking's Global Ranking of Sport Science Schools and Departments*, 2021) reveals similar results. While two universities offer an introductory course on statistics (being a relevant part of data science), one of the examined curricula does not even include such a basic course. Furthermore, none of the examined curricula included courses on modelling or data processing, which would be necessary to develop actionable knowledge in data science (De Veaux et al, 2017).

4. Closing the Gap?

At present, coaches and sport scientists besides educating themselves through self-study (e.g., using offers on the internet) can also enrol into extra-curricular offers of universities (e.g. courses such as "Introduction to Programming"). However, these choices often have the downside of being too detailed and targeted at other audiences e.g. computer science students. Recently, the possibility of targeted continuing education courses has emerged. These courses often offer "graduate certificates" and follow a strict curriculum (Victoria University, 2021).

Another potential solution are projects aiming to promote and advance programming education targeting broader (non-computer science) audiences that might prove valuable if integrated, for example, as extra-curricular activities into sport science curricula. One example of such a project is Codability (CodeAbility Austria, 2021) aiming to provide programming courses to broad audiences.

Yet another solution might be a shift in existing curricula towards the integration of these topics into existing courses. For sport scientists, data science competences might be considered as transversal skills. For example, the ongoing ATSSTEM project aims at the coherent development of transversal skills in an integrated STEM (science, technology, engineering and mathematics) curriculum. It provides educators with the formative digital assessment of transversal skills as learners develop real-world and authentic STEM competences (Costello et al., 2021). Similarly, the use of ICT tools can support integrating data science topics in sport science curricula.

5. Conclusions

The rising demands on sport scientists with respect to data analysis should also be reflected in their professional education and development. In order to not be left behind, the education of (future) sport scientist needs to improve with regard to data science and related topics. Additionally, the contents of continuing education programs should not remain on the level of learning to operate specific tools, but has to aim to develop an understanding for general concepts in data science, such as "computational thinking". Only in this way, learners can be supported to develop transferable skills that can be adapted to the opportunities and challenges emerging with the continuing evolution of technical possibilities in data capturing and processing.

References

1 CodeAbility Austria. (2021). https://codeability.uibk.ac.at/index-en.html

2 Costello, E., Aral, N., Hartell, E., Lennholm, H., Kirwan, C., Girme, P., & Brown, M. (2021). Is it in the bin? Seeking authentic assessment in STEM : ATSSTEM. PATT38 Book of Abstracts, 31–31

3 De Veaux, R.D., Agarwal, M., Averett, M., Baumer, B.S., Bray, A., Bressoud, T.C., Bryant, L., Cheng, L.Z., Francis, A., Gould, R. & Kim, A.Y. (2017). Curriculum guidelines for undergraduate programs in data science. Annual Review of Statistics and Its Application, 4, pp.15-30.

4 Dhar, V. (2013). Data science and prediction. Communications of the ACM, 56(12), 64–73. https://doi.org/10.1145/2500499

5 Dobiasch, M., & Oppl, S. (2020). Modellabstraktionen und deren Wirkung im Sport – Ein Ausbildungsproblem? Wirtschaftsinformatik & Management, 12(5), 362--368. https://doi.org/10.1365/s35764-020-00282-y

6 ShanghaiRanking's Global Ranking of Sport Science Schools and Departments. (2021). https://www.shanghairanking.com/rankings/grsssd/2020

7 Victoria University. (2021). Graduate Certificate in Data Analytics for Sport Performance. Victoria University | Melbourne Australia. https://www.vu.edu.au/courses/graduate-certificate-in-data-analytics-for-sport-performance-stsp