

## Guest editorial: Special issue on sports analytics

Ulf Brefeld<sup>1</sup> · Albrecht Zimmermann<sup>2</sup>

Received: 29 June 2017 / Accepted: 7 July 2017 / Published online: 25 July 2017  
© The Author(s) 2017

This special issue is dedicated to recent advances in sports analytics. The field got its start as a somewhat exotic area in the statistics community ([MathSport International 2017](#)) and eventually attracted interest from sports management professionals (as immortalized by the book “Moneyball”; [Lewis 2004](#)), culminating in the “MIT Sloan Sports Conference” ([MIT Sloan School of Management 2017](#)).

As in most other areas of society, increasing amounts of data are being collected in all kinds of sports, and *automated* data analysis has become an important and rapidly developing field. Hence, more and more researchers from data mining (and machine learning) communities have joined the fray and specialised workshop series such as “Machine Learning and Data Mining for Sports Analytics” at ECML/PKDD ([Davis et al. 2013](#)) or “Large-Scale Sports Analytics” at KDD ([Lucey et al. 2016](#)) attest the increasing popularity of intelligent data analysis for sports analytics.

Sports analytics resides at a very attractive intersection of concrete application problems (both in health care/medical/sports research and the various industries connected to professional sports), significant commercial interest, large data collections, and wide-spread interest by the general public. Similar to early artificial intelligence research on games, sports data offer themselves as an interesting test bed for data mining researchers. In contrast to games—most of which have by now been solved—the

---

Responsible editor: Ulf Brefeld and Albrecht Zimmermann.

---

✉ Albrecht Zimmermann  
albrecht.zimmermann@unicaen.fr  
Ulf Brefeld  
brefeld@leuphana.de

<sup>1</sup> Machine Learning, Leuphana University, Lüneburg, Germany

<sup>2</sup> UNICAEN, ENSICAEN, CNRS, GREYC, Normandie Univ, 14000 Caen, France

universe of sports data is not discrete, and the elements of human individuality and human error have more influence, increasing the challenge. Yet at the same time many sports are constrained by rules and the place (track, field, climbing wall etc.) where the action takes place, offering repeatability that other real-life settings lack. Finally, sports data are being systematically collected, without the gaps (or mis-assigned values) that occur in sensor or administrative data.

Today, many different strategies, methods, and techniques have been studied in the context of sports analytics, depending on the type of sports, the data, and the goals of the analysis. Consequently, despite sports analytics being a rather novel field, there already exists a very diverse set of research questions, approaches, and data sources in the literature. Hence, the collection of papers presented in this special issue is neither comprehensive nor anticipates all future trends, but we hope it will provide a starting point for entering this exciting area.

As we expected, the special issue received much interest, attracting twenty eight submissions, of which ten made the final cut. The material included in this collection covers sports as diverse as baseball (Bendtsen et al.), basketball (van Bommel et al., Vinue et al.), beach volleyball (Kautz et al.), cricket (Salman et al.), ice hockey (Schulte et al.), rock climbing (Herault et al.), soccer (Andrienko et al., Kostakis et al.), and speed skating (Knobbe et al.). The reported analyses are performed both on an aggregated (match actions, referee bias etc) and on individual levels (player movements, player diets etc).

The contributions in this special issue challenge the standard computation model of learning from independently drawn instances by focusing on spatio-temporal player trajectories, regularly conducted physiological measurements, or player career data. Technically, the analyses are carried out using a variety of approaches, such as deep learning, Bayesian networks, or archetypal analysis. In contrast to typical data mining papers, in many cases the described work is not limited to the data analysis technique but involves the entire KDD pipeline from data preparation (or even data acquisition) via transformation, analysis, up to the interpretation of derived results.

Some articles extend contributed papers of the afore-mentioned workshops, and the Sloan Sports Analytics Conference, but this special issue also contains novel contributions. We thus believe that it provides a timely snap-shot of the state-of-the-art in this exciting field, covering a wide range of methods and sports.

We would like to thank everybody who contributed to the special issue, the authors for their valuable contributions, the guest editorial board for delivering high-quality reviews, Melissa Fearon and the Springer team for making it happen and, last but not least, Juffi Fürnkranz for pushing the idea of a special issue forward. Enjoy reading,

Ulf Brefeld and Albrecht Zimmermann,  
29/06/2017

## References

- Davis J, van Haaren J, Kaytoue M, Zimmermann A (2013) Machine learning and data mining for sports analytics. <https://dtai.cs.kuleuven.be/events/MLSA17/>
- Lewis M (2004) Moneyball: the art of winning an unfair game. WW Norton & Company, New York
- Lucey P, Morgan S, Wiens J, Yue Y (2016) KDD workshop on large-scale sports analytics. <http://www.large-scale-sports-analytics.org/>
- MathSport International (2017). <http://www.mathsportinternational.com>
- MIT Sloan School of Management (2017) MIT sloan sports analytics conference. <http://www.sloansportsconference.com>