

特集：トップスポーツを支援する情報システム・分析技術および
大学体育の教育活動支援

スポーツアナリティクスにおけるデータとAI活用

谷岡 広樹*

Data and Artificial Intelligence Utilization Required for Sports Analytics

Hiroki TANIOKA*

Tokyo 2020 Olympic Games has been postponed until 2021. Most of the 33 sports still planned for the Olympic Games in 2021 will use data. The sports data gathered using various method is analyzed by experts. The experts also called sports data analysts have been developed various systems and methods using the sports data and Artificial Intelligence to improve competition results. In this paper, data collections in some sports are explained, and the trends in research on the use of data analysis models in these fields, in particular, some studies using machine learning and our research are introduced. Additionally, future issues in this field are summarized.

キーワード：スポーツアナリティクス，データサイエンス，Artificial Intelligence (AI)，機械学習

1. はじめに

本来ならばオリンピックイヤーであった2020年⁽¹⁾は、スポーツのあらゆる種目でデータの活用が進んでおり、かつては経験・勘・根性といったものが主役であった姿からは一変し、データを用いた科学的な分析や深層学習などの機械学習を用いたデータ分析が行われている⁽²⁾⁽³⁾。スポーツアナリティクス (Sports Analytics: SA) と呼ばれるデータ分析は、統計学を背景とした手法が用いられ、スポーツ統計⁽⁴⁾とも呼ばれており、データサイエンス (Data Science: DS) とも密接な関わりを持つ。

このようなスポーツにおけるデータ活用は、1960年代の米国のプロ野球リーグ (Major League Baseball: 以下, MLB) を舞台に展開された映画「マネー・ボール」の題材となったセイバーメトリクス (Society for American Baseball Research Metrics: SABR metrics) が1970年代に提唱されて以降、急速に広まったが、そのデータ活用の歴史は、1861年にHenry

Chadwick が Beadle's Dime Base Ball Player⁽⁵⁾ において、野球のプレーを分析するためには、統一された方法でデータを収集しなければならないことを述べたことから始まる。その後の1885年にThe Lawn Tennis Manual⁽⁶⁾では、テニスプレイヤーのパフォーマンス分析を行った。1900年代に入り、アメリカンフットボール、ラグビー、サッカーでも、試合から客観的なデータを収集し、記事として掲載されるようになっていった。1922年にはThe Science of Baseball⁽⁷⁾で野球データの記法が確立したといわれている。1930年代以降には、欧州のプロサッカーチームや米国のMLB、アメリカンフットボールリーグ (American Football League: AFL) ナショナルホッケーリーグ (National Hockey League: NHL)、ナショナルバスケットボールリーグ (National Basketball Association: NBA) でアナリストを採用するようになった。このように、スポーツアナリティクスの歴史は、スポーツの商業面での貢献を目的として発展し、2000年以降は、スポーツデータ分析を専業とする企業^{(8)~(12)}も国内

* 徳島大学情報センター (Center for Administration of Information Technology, Tokushima University)

外に存在する状況となっている。

オリンピック競技のように、商業面よりも競技成績に重きを置く競技においても、データ分析やAIを活用した戦略や戦術の変更、選手の起用などが盛んに行われるようになってきている⁽¹³⁾⁽¹⁴⁾。このとき、対戦競技においては戦術論とは別に、各選手の能力向上のためのトレーニングにデータを活用することも盛んに行われるようになってきた。野球であれば投球フォーム、打撃フォームなどの改善に選手の体の動きをデータ化したものが利用される場面も増加している。また、水泳や体操のような個人競技や演技種目においても、同様のデータ分析の手法を用いることによる成績の向上が期待されている⁽¹⁵⁾⁽¹⁶⁾。

本稿では、2章でスポーツにおけるデータ収集の方法について紹介したうえで、3章でデータ分析手法に関する研究動向、4章でデータ分析の活用事例とデータ活用に関する課題などについてまとめる。

2. スポーツにおけるデータ収集の方法

本章では、スポーツデータの収集方法に関する状況について概説する。スポーツ競技で競技成績を向上するためには、表1に示すデータのうち、情報システムによって収集・分析が可能なものを対象とし、選手個人のデータ、対戦競技の場合はその対戦相手のデータ、球技の場合は球に関するデータ、団体競技の場合は味方チームの選手の位置データ、相手チームの選手の位置データなどの収集が必要となる。

2.1 IoT センサ・GPS データ

競技中の戦術変更やトレーニングへのフィードバックのために必要となるデータを、どのように収集するかについては、各競技においてさまざまな工夫がなされている。本節では、モノのインターネット (Internet of Things: IoT) センサや全球測位システム (Global Positioning System: GPS) データを用いて収集可能なデータを、トレーニングに活用する事例について紹介する。

2.1.1 IoT センサ

IoT センサは、ウェアラブルのものと、観測対象から離れた位置に設置されるものに分類できる。ウェア

ラブルセンサには、次項で紹介するGPS データを検出可能なもの⁽¹⁷⁾のほか、加速度、脈拍、酸素濃度、筋電位、さらには血糖値も測れるデバイス⁽¹⁸⁾⁽¹⁹⁾が登場してきており、トレーニングや試合の前後、トレーニング中に加えて、規定が許せば試合中のデータまで得ることが可能となり、選手の体調管理はもちろんトレーニングメニューや戦術に活かすことが可能である。

観測対象から離れた位置に設置されるIoT センサには、Light Detection and Ranging (LIDAR) 技術など、3D レーザーセンサを用いて選手の姿勢データや位置データを正確に把握するもの⁽²⁰⁾や、バットやボールにセンサを内蔵するもの、競技場自体に圧力センサや超音波センサを配置することで、さまざまな競技データを収集する試み⁽²¹⁾もある。これらは大会規定が許せば、競技中にリアルタイムにフィードバックするシステムによってデータ分析を行うことができるが、規

表1 データの分類

カテゴリ	対象	変数例
人物	競技者	身長, 体重, 体温, 脈拍, 姿勢, 座標, 方向, 速度
	対戦者	身長, 体重, 体温, 脈拍, 姿勢, 座標, 方向, 速度
	審判	座標, 方向, 速度
オブジェクト	ボール	サイズ, 座標, 方向, 速度
	バット	サイズ, 座標, 方向, 速度
	ラケット	サイズ, 座標, 方向, 速度
	ネット	サイズ, 座標
	ゴール	サイズ, 座標
	器具	サイズ, 材質
	床面	サイズ, 材質, 床反力
	水	水量, 深さ, 温度, 水流
	衣類	材質, 重量, 形状
環境	気象条件	屋内外, 温度, 湿度, 気圧, 風向, 風速
	建物	サイズ, 材質
	観客席	サイズ, 座席数, 距離

定により採用不可な場合もある。

2.1.2 モーションキャプチャデータ

選手の姿勢や体重移動など、選手の動きに着目したデータを収集する方法の一つに、モーションキャプチャがある。この方法では、全身にマーカーを付けてマーカーの座標を測定するモーションキャプチャ技術⁽²²⁾のほかに、Microsoft社のKinectやLIDAR技術を採用したレーザーセンサなどを用いたマーカーレスの方法⁽²³⁾や、次節で詳しく説明する深層学習を用いた手法も利用可能である。

2.1.3 GPS データ

集団競技の場合は、骨格の検出よりも人物やオブジェクトの座標データとトラッキングデータが重要となるケースがある。このような場合は、Catapult社⁽²⁴⁾などのGPSセンサを選手に着用させることで正確な位置データやトラッキングデータを取得する方法が考えられ、2015年から国際サッカー連盟(以下、FIFAと呼ぶ)、2016年からJリーグでも試合での着用在認められている。

2.2 動画データ

マーカーレスのモーションキャプチャ技術を実現するには、KinectやLIDAR技術を採用したレーザーセンサなどの特別な機材を用いる方法のほかに、画像からOpenPose⁽²⁵⁾やPoseNet⁽²⁶⁾などの深層学習を用いて特徴点を抽出する方法⁽²⁷⁾がある。これらの技術は、画像の中の人物の骨格を検出し、高い精度で各部位のベクトルから姿勢データを得ることができるため、特別な機材を用いずに選手の姿勢や動きを認識することが可能となる。

集団競技の場合に重要となる座標データとトラッキ



図1 YOLOv3で抽出された人物候補領域
(図は文献(31)より引用)

ングデータを得るために、GPSデータを用いずに、カメラ動画から検出する方法もある。ChyronHego社のTRACAB⁽²⁸⁾は、FIFAやJリーグでも採用されている光学式トラッキングシステムであり、複数のカメラを用いて選手やボールの座標を正確にデータ化することが可能である。このほか、単眼カメラ1台の映像からYOLOv3などの深層学習と画像解析技術を用いて図1のように人物の座標やトラッキングデータを抽出する手法^{(29)~(31)}も提案されており、この手法を用いると複数の機材を会場に設置することや、対戦相手に機器着用の同意を得る必要がない。

3. スポーツにおけるデータ分析

3.1 スポーツデータと統計データ

スポーツデータは、IoTセンサやビデオカメラなどで得られた未加工の生データのままで、データ分析が困難である。そのため、この生データから統計データ(Statistics: Stats)の形にフィルタリングや加工を施す必要がある。Shihは、動画データからStatsを抽出する場合の方法について詳しくまとめている⁽³²⁾。このなかで、生データには競技そのものの情報以外に、観客や動物乱入などのシーンが含まれているため、不要な情報の除去も必要であると述べている。

3.2 スポーツデータとデータ分析

Statsを得た後、データ分析者は目的を達成するための課題を競技やチームごとに定義し、統計的な分析を行う。加藤はJリーグでのサッカーチームにおけるチーム強化のためのゲーム分析について紹介⁽³³⁾している。一方、課題自体が曖昧な場合もある。このような場合は、あらゆる角度から統計的に予測モデルを構築して回帰分析を行う手法や、データマイニングのツールを用いる手法⁽³⁴⁾⁽³⁵⁾が考えられる。

3.3 スポーツデータと人工知能

急速に発達した深層学習の技術を用いた人工知能によるデータ分析も進みつつある^{(13)~(16)}。オブジェクト検出等の分野でSotA(State of the Art)を達成している深層学習⁽³⁶⁾を用いたデータ分析は、従来法と比較して高速かつ高精度にStatsを得ることがで

き、再現率 (Recall)、適合率 (Precision)、ROC 曲線 (Receiver Operating Characteristic Curve) や AUC (Area Under the Curve) などの指標で比較して、高い予測精度でモデルを推定することが可能である。

4. スポーツにおけるデータ活用とその課題

この章では、スポーツにおけるデータ活用の事例を紹介する。4.1 節では選手の能力評価、4.2 節では集団競技における戦術分析について紹介し、4.3 節では、データの品質に関する課題、4.4 節では権利問題について述べる。

4.1 ボールトラッキングデータによる能力評価

文献 (37) は投球トラッキングデータと SABR metrics を用いて MLB の投手のパフォーマンスの予測を試みている。文献 (38) では、打球トラッキングデータを用いてロジスティック回帰分析を行うことにより、日本野球機構 (NPB) の打者の評価を試みている。

4.2 選手トラッキングデータによる戦術分析

文献 (39) は、NBA の選手とボールのトラッキングデータを用いて戦術分析を行い、戦術選択を正しく学習できるシステムとそれを用いたトレーニングの方法を提案している。文献 (40) では、欧州のプロサッカーリーグの Stats に基づいて、期待得点 (expected Goals: xG) を定義している。

4.3 データの品質に関する課題

スポーツにおけるデータ活用のためには、データそのものの品質の問題について注意する必要がある。選手やチームの状況を正確に把握し、統計的に意味のある予測モデルを構築するためには、十分なデータをバランスよく収集する必要がある。収集したデータが一部の選手や特定の条件下のものしかないような不均衡データであった場合、正しく統計分析を行うことは困難である。

データの取得条件に、環境情報が含まれることもある。観測対象の競技場の形状はもちろん、観測対象と機材の距離や角度、気象条件や日照状態の変化によっ

ても、選手や競技内容とは無関係にばらつきが現れる。このような環境の動的な変化に対して対策を講じる必要がある。

運良く理想的なデータが得られた場合、よりよい分析結果や予測モデルを得るために、われわれ研究者は複雑な予測モデルや人工知能の技術を採用することがあるだろう。このとき、考慮すべき問題の一つに、モデルの説明可能性が挙げられる。人工知能のシステムが採用するほとんどの機械学習アルゴリズムでは、学習結果として得られるモデルについて、解釈不可能な場合が少なくない。その結果、構築された予測モデルはブラックボックスとなり、なぜそのような予測が得られるかについて説明できないという事態に陥る。この問題を回避するため、回帰モデルを利用する場合は、線形回帰モデルやロジスティック回帰モデルを利用すること、機械学習を利用する場合は、決定木やランダムフォレストのような説明変数の寄与度を算出可能なアルゴリズムを用いるとよい。

4.4 スポーツデータと権利

スポーツデータは、大会の主催者や放映権を持つ事業者などに権利がある。そのため、自由にデータ分析を行うためには、データの権利者から利用許諾を得るか、自ら収集するしかない。スポーツデータ解析コンペティション⁽⁴⁾ 参加を対象に研究目的で公開されるデータ、StatsBomb Open Data⁽⁴¹⁾、NBA Stats⁽⁴²⁾ などが公開されているが、利用範囲に注意が必要である。

5. おわりに

本稿では、スポーツにおけるデータ活用の目的、データ収集の方法、データ分析の手法、人工知能や機械学習の活用方法の研究動向や今後の課題について概説した。

近年の深層学習を用いた機械学習の発展は、スポーツデータの解析や予測モデルの構築に大きな影響を与えつつある。しかしながら、4 章で述べたように、データの品質の問題や、権利の問題から、自由に利用できるデータは十分とはいえず、競技成績の向上や選手の能力向上に資するシステムを構築し、活用するためには、より可用性の高いシステムの開発と、さらな

るルール改正や実践事例の蓄積が必要である。

謝辞

本稿の一部は JSPS 科研費 JP18H03344 の助成を受けたものです。

参考文献

- (1) 東京 2020 オリンピック : <https://tokyo2020.org/>(参照 2020.4.26)
- (2) 国立スポーツ科学センター年報 2018 : <https://www.jpnsport.go.jp/jiss/gaiyou/jigyou/houkoku/nenpo/tabid/260/Default.aspx> (参照 2020.4.26)
- (3) 相原伸平, 杉山恵玲奈, 澤田みのり, 松本 実, 伊藤浩志 : “競技スポーツの実践現場における ICT 活用”, 電子情報通信学会通信ソサイエティマガジン, Vol. 12, No. 2, pp. 98–104 (2018)
- (4) 日本統計学会スポーツ統計分科会 : <https://estat.sci.kagoshima-u.ac.jp/sports/>(参照 2020.4.26)
- (5) Chadwick, H.: “Beadle’s Dime Base-Ball Player: A Compendium of the Game, Comprising Elementary Instructions of this American Game of Ball, Together with the Revised Rules and Regulations for 1860, Rules for the Formation of Clubs, Names of the Officers and Delegates to the General Convention, & c.”, Irwin P. Beadle & Co. (1860)
- (6) Chadwick, H.: “The Lawn Tennis Manual”, A. G. Spalding & Bros, New York (1885)
- (7) Byrd, D.: “The Science of Baseball: A Text-Book of “inside” Baseball Completely Covering Every Department and Phase of Baseball: How to Play and Coach the Game”, T. E. Wilson (1922)
- (8) STATS LLC: <https://www.stats.com/>(参照 2020.4.26)
- (9) Hudl: <https://www.hudl.com/>(参照 2020.4.26)
- (10) Gracenote Sports: <https://www.gracenote.com/>(参照 2020.4.26)
- (11) Opta Index: <https://www.optasports.com/>(参照 2020.4.26)
- (12) Data Stadium: <https://www.datastadium.co.jp/>(参照 2020.4.26)
- (13) 清水千弘, 清田陽司 : “スポーツ競技と AI”, 人工知能学会誌, Vol. 34, No. 4, pp. 492–496 (2019)
- (14) 加藤健太 : “スポーツデータの概要と AI 活用の可能性”, 人工知能学会誌, Vol. 34, No. 4, pp. 539–544 (2019)
- (15) 青木義満 : “スポーツ競技を対象とした画像 AI 技術”, 人工知能学会誌, Vol. 34, No. 4, pp. 497–502 (2019)
- (16) 金澤裕治, 榊井昇一, 矢吹彰彦, 佐々木和雄 : “体操自動採点に向けた 3D センシングと技認識の AI 技術”, 人工知能学会誌, Vol. 34, No. 4, pp. 531–538 (2019)
- (17) 向本敬洋, 伊藤雅充, 河野徳良, 野村一路, 西條修光 : “GPS 機器を利用した大学男子サッカー選手における各ポジションの Time-motion 分析”, コーチング学研究, Vol. 27, No. 2, pp. 215–223 (2014)
- (18) 馬込卓弥 : “IoT/センサ技術を用いたスポーツ分析とスポーツ医学への展開”, 電子情報通信学会通信ソサイエティマガジン, Vol. 12, No. 2, pp. 112–116 (2018)
- (19) Harsha Vardhan Goud, P. S., Mohana Roopa, Y. and Padmaja, B.: “Player performance analysis in sports: With fusion of machine learning and wearable technology”, 3rd International Conference on Computing Methodologies and Communication (ICCMC), pp. 600–603 (2019)
- (20) 佐々木和雄, 榊井昇一, 手塚耕一 : “アスリートの動きをリアルタイムに数値化する 3D センシング技術 (特集 スポーツ ICT—スポーツデジタルプラットフォームを支えるテクノロジー—”, Fujitsu, Vol. 69, No. 2, pp. 13–20 (2018)
- (21) 長岡大志, 荻野尚哉, 坪井一洋, 野田茂穂, 姫野龍太郎 : “加速度センサを用いた飛翔中の野球ボールの回転特性計測”, 日本機械学会論文集, Vol. 85, No. 867, pp. 1–15 (2019)
- (22) 角川隆明, 萬久博敏, 荻田 太 : “クロールにおける泳速度の変化とストロークパラメーターの関係—圧力分布計測と水中モーションキャプチャを用いた分析—”, Vol. 64, No. 1, pp. 385–400 (2019)
- (23) 春名弘一, 昆 恵介, 稲垣 潤, 佐藤洋一郎 : “マーカーレスモーションキャプチャによる三次元動作解析の応用例”, 日本義肢装具学会, Vol. 35, No. 1, pp. 17–23 (2019)
- (24) Catapult: <https://www.catapultsports.com/>(参照 2020.4.26)
- (25) Cao, Z., Hidalgo Martinez, G., Simon, T. et al.: “OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields”, IEEE Transactions on Pattern Analysis and Machine Intelligence, p. 1 (2019)
- (26) Kendall, A., Grimes, M. and Cipolla, R.: “PoseNet: A Convolutional Network for Real-Time 6-DOF Camera Relocalization”, Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), IEEE Computer Society, No. 9, pp. 2938–2946 (2015)
- (27) 姜 文淵, 山本雄平, 田中成典, 中村健二, 田中ちひろ : “単視点多眼によるアメリカンフットボールプレイヤー

- の識別と位置特定に関する研究”, 写真測量とリモートセンシング, Vol. 57, No. 5, pp. 198-216 (2018)
- (28) TRACAB: <https://chyronhgo.com/products/sports-tracking/tracab-optical-tracking/>(参照 2020.4.26)
- (29) 大内一成, 小林大祐, 中洲俊信, 青木義満: “ラグビー映像解析システムの開発”, 電子情報通信学会論文誌 B, Vol. J100-B, No. 12, pp. 941-951 (2017)
- (30) Xu, J. and Tasaka, K.: “Keep your eye on the ball: Detection of kicking motions in multi-view 4K soccer videos”, ITE Transactions on Media Technology and Applications, Vol. 8, No. 2, pp. 81-88 (2020)
- (31) Karungaru, S. G., Matsuura, K., Tanioka, H. et al.: “Towards drone-video player detection and tracking for soccer strategy analysis”, Proceedings of the 6th IEEJ international workshop on Sensing, Actuation, Motion Control, and Optimization, Vol. TT-1, pp. 1-6 (2020)
- (32) Shih, H.-C.: “A survey of content-aware video analysis for sports”, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 28, No. 5, pp. 1212-1231 (2018)
- (33) 加藤健太: “サッカーにおけるデータ分析とチーム強化”, 電子情報通信学会通信ソサイエティマガジン, Vol. 10, No. 1, pp. 29-34 (2016)
- (34) Parmezan Bonidia, R., Duilio Brancher, J. and Marques Busto, R.: “Data mining in sports: A systematic review”, IEEE Latin America Transactions, Vol. 16, No. 1, pp. 232-239 (2018)
- (35) Apostolou, K. and Tjortjis, C.: “Sports Analytics algorithms for performance prediction”, 10th International Conference on Information, Intelligence, Systems and Applications (IISA), pp. 1-4 (2019)
- (36) Wang, H. and Zheng, X.: “Survey of deep learning based object detection”, Proceedings of the 2nd International Conference on Big Data Technologies, Association for Computing Machinery, No. 5, pp. 149-153 (2019)
- (37) 松木拓弥, 鈴木秀男: “トラッキングデータとセイバーメトリクス指標による MLB 投手の能力評価モデルの構築”, 電子情報通信学会通信ソサイエティマガジン, Vol. 12, No. 2, pp. 117-125 (2018)
- (38) Sakai, M., Tanioka, H., Matsuura, K. et al.: “Evaluating hitting skills of npb players with logistic regression analysis”, CSCE'18 Proceedings, pp. 313-319 (2018)
- (39) 箭野 柊, 松浦健二, 谷岡広樹, カルンガル・ギディンシ・スティフィン, 幸田尚也, 後藤田中, 和田智仁: “集団対戦型フィールドスポーツの戦術適用判断の支援環境—バスケットボールのオフENSES基本戦術—”, 情報処理学会論文誌, Vol. 61, No. 3, pp. 657-666 (2020)
- (40) Rathke, A.: “An examination of expected goals and shot efficiency in soccer”, Journal of Human Sport and Exercise, Vol. 12, Vol. 2, pp. s514-s529 (2017)
- (41) StatsBomb Open Data: <https://github.com/statsbomb/open-data> (参照 2020.4.26)
- (42) NBA Stats: <https://stats.nba.com/>(参照 2020.4.26)

著者紹介



谷岡 広樹

1997年千葉大学工学部卒。2008年信州大学大学院総合工学系研究科博士課程修了。博士(工学)。1997年ジャストシステム, 2011年古河インフォメーション・テクノロジー, 2014年ワークスアプリケーションズ, 自然言語処理, 情報検索, 機械学習に関する研究開発に従事。2016年より現職。教育システム情報学会2018年度研究会優秀賞。教育システム情報学会, 情報処理学会, 人工知能学会, IEEE, ACM等会員。