

# 機械学習を用いたバレーボールの勝敗予測

2020SE018 蔭山宗佑

指導教員：野呂昌満

## 1 はじめに

近年、スポーツの中で戦略とデータ処理の関係は重要視されている。その中でも本研究では、データバレーという言葉が存在するほど、データ処理と密接な関係にある、バレーボールに着目する。

競技力向上のため [1]、健康管理のため [2] というようなデータ処理の活用がある中、本研究では、まだ先行研究が少ない、勝敗予測について研究を行う。勝敗予測を行う上での問題点として、どのスポーツでも共通して、勝敗に影響を与える要因が数多く存在していることが挙げられる。

著者は、バレーボールのような競技において、特に趨勢が勝敗を左右する大きな要因であると考ええる。本研究では、その趨勢が、試合結果を左右するといった仮説を立てた後、その仮説の検証を行うことを目的とする。

本研究の課題は、目的を達成するための実験の精度を向上させることである。実験の精度を向上させるための技術課題は、RQ1：特徴量の決定、RQ2：モデルの作成、RQ3：作成したモデルの検証、評価である。

## 2 関連研究

米沢ら [3] の研究では、バレーボールゲームのセットの序盤・中盤・終盤の得点差から勝敗予測を行った。米沢らは、1999 年度よりバレーボールに導入されたラリーポイント制が、サイドアウト制からの移行によってゲームの内容に様々な影響を及ぼしていると考えている。米沢らは、そのラリーポイント制のゲーム戦術に示唆を与えることを目的として研究を行った。結果としては、得点差と勝率にはかなり高い相関関係があり、ゲームの勝敗を予測することが十分可能であると述べている。

## 3 課題解決のアプローチ

### 3.1 システムの設計

本研究では、趨勢は連続得点や連続失点、エースのポジションによって左右されると定義し、その趨勢を数値化したものをもとに勝敗予測を行う。勝敗予測は、二値分類問題の実験ではよく使われる手法であり、他の統計手法と比較して、理解するのが簡単であるという理由から、ロジスティック回帰分析を用いて行う。システムにおける処理手順は図 1 に、趨勢の数値化の方法は表 1 に示す。

エースが前衛にいる時は、趨勢の移り変わりが激しいと考え、前衛にいる時に連続得点をした場合、趨勢の値を +2 とする。連続失点をした場合には、趨勢の値を -2 とする。本研究でいうエースは、出場セットが多い、アタックの打数が多い、アタック決定率が高いという 3 つの条件に当てはまる選手とする。エースが後衛にいた場合は、それ

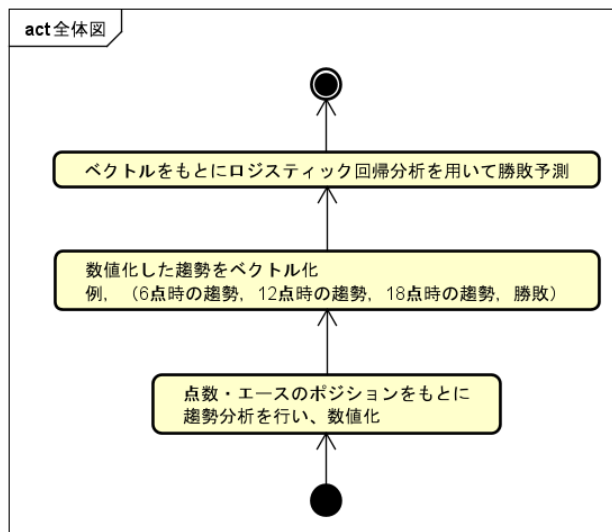


図 1 システムにおける処理手順

表 1 趨勢の数値化

エースの前後	連続	趨勢
前	得点	+ 2
前	失点	- 2
後ろ	得点	+ 1
後ろ	失点	- 1

ぞれの値を +1, -1 とし、エースが前衛にいる時と後衛にいる時での差別化を図った。負の値になることを防ぐため、趨勢の最初の値を 10 とし、数値化した趨勢の値を、点数毎に累計的に扱い、配列にしていく。

このように数値化し、配列にした趨勢をもとに、ロジスティック回帰分析を行うためのデータとして、以下のベクトルを用意する。

(6 点時の趨勢, 12 点時の趨勢, 18 点時の趨勢, 勝敗)

配列となった趨勢をもとに、6 点時、12 点時、18 点時の趨勢と勝敗をベクトル化する。勝敗については、対象チームが勝利の時は 1、敗北の時は 0 とした。このベクトルの 6 点時、12 点時、18 点時の趨勢を説明変数とし、勝敗を目的変数としたロジスティック回帰分析を行っていく。

### 3.2 対象データ

本研究では、Web サイトを使用して、趨勢が左右されると考える、点数、エースのポジションの 2 つの特徴量を抽出する。選手の調子などとは違い、日によってばらつきが存在しない点数や、連続得失点時のエースのポジションで

趨勢に与える影響が異なると考えたので、この特徴量を選択した。

## 4 実装環境

本研究では実装環境として、コードの逐次実行が可能であること、記録や報告目的で使いやすいこと、分析で用いるデータをクラウドに保存する必要がないという理由から、Jupyter Notebook を用いる。Jupyter Notebook は、インタラクティブにプログラムを実行できる web ベースの Python の環境のひとつである。

### 4.1 データセット

使用するデータは過去 3 年分の試合結果とする。チームの方針によってエースを一人に絞らないチームもあるので、本研究では、愛知県稲沢市に本拠地を置く、ウルフドッグス名古屋チームに絞ってデータセットを作成する。データ収集には、Livesport\*<sup>1</sup>、バレーボール v リーグオフィシャルサイト\*<sup>2</sup>を使用する。作成したデータセットは全部で 367 行×4 列となった。

### 4.2 実験及び評価

作成したデータセットをもとに実験を行った結果を、以下の表 2 に示す。

表 2 実験結果

指標	値
accuracy	0.8378378378378378
precision	0.9056603773584906
recall	0.8727272727272727
f1 score	0.8888888888888888
auc	0.8899521531100478

実験の結果から、auc の値はおよそ 0.89 と、かなり高い値を得ることができた。他にも、accuracy, precision, recall, f1 score の値もすべて 80% を超える結果となった。

## 5 考察

### 5.1 仮説の妥当性

4.2 節の結果から、バレーボールの試合結果には趨勢が関係しているという仮説の妥当性を考察する。

本研究では、趨勢を点数・エースのポジションという 2 つの特徴量から判断した。点数については、特に連続得点・連続失点に着目し、この特徴量は、趨勢を判断するのに最も容易な特徴量であると考えられる。選手成績なども同様に参考になると考えるが、その日の選手の調子などで、日によってばらつきが存在する。得点に着目することで、そのばらつきによる誤差を少なくすることができたと考える。

\*<sup>1</sup> <https://www.livesport.com/jp/team/nagoya/jTKINSpR/>

\*<sup>2</sup> <https://www.vleague.jp/record/match/>

以上のことから、バレーボールにおける試合結果は、趨勢によって左右されるという仮説の検証が行えたといえる。

本研究を利用すると想定する場面として、テクニカルタイムアウトが考えられる。バレーボールの国際大会には、テクニカルタイムアウトという制度が与えられており、これは 1~4 セット目で勝っているチームが 8 点もしくは 16 点になった時、設けられる休憩時間である。この時間で戦術を推測するために、8 点時、16 点時の趨勢を用いた勝敗予測も行った。auc が 0.8 を超える実験結果となり、このことから戦術の推測に利用できるのではないかと考える。

### 5.2 今後の課題

今後の課題として、特徴量の調整、データセットの増加、RNN による勝敗予測が挙げられる。

本研究では、取り扱った特徴量が 2 つであったが、この特徴量を追加したり、変更したりすることで実験の結果がさらによくなるのではないかと考える。

また本研究では、ロジスティック回帰分析を用いた実験を行ったが、本研究で用いた特徴量は、時系列データとして取得されるので、RNN を用いた実験を行うことによって、さらに時系列を考慮した、精度の高い結果が得られると考える。

## 6 おわりに

どのスポーツでも重要とされている戦略とデータ処理の関係であるが、今回は特に注目されているバレーボールにおいて、先行研究が少ない勝敗予測について研究を行った。本研究では、バレーボールの試合結果には、趨勢が関係しているという仮説を立て、この仮説を検証することで問題の解決に臨んだ。数値化した趨勢をベクトル化し、このベクトルをもとに勝敗予測を行い、得られた結果から、システムの検証、評価を行った。結果は、どの評価指標も、高い数値を得ることができた。この結果から、仮説の検証を行うことができた。今後の課題としては、特徴量の調整、データセットの増加、RNN による勝敗予測が挙げられる。

## 参考文献

- [1] 江崎 修央, 重永 貴博, 宮地 力, “バレーボールゲーム分析システム TOUCH VOLLEY における戦術支援機能とデータ分析機能の実装,” バレーボール研究, 6 巻 1 号, p.29-34 (2004.5).
- [2] 横尾 優美, 武本 歩未, 大塚 美智子, “肩関節をサポートするバレーボール用インナーウェアのパターン提案,” 日本女子大学大学院紀要. 家政学研究科・人間生活学研究科 (2021).
- [3] 米沢 利広, 松本 勇治, 俵 尚申, “バレーボールゲームにおける勝敗の予測-大学女子バレーボールチームについて-,” バレーボール研究, 2 巻 1 号, p.29-35 (2002.5).