

機械学習によるサッカーの戦術分析

Y210121 奥村星哉

指導教員 三好 力 教授

1, はじめに

データ解析の進展により、サッカーではトラッキングデータを活用した選手やチームの評価が進んでいる。しかし、試合状況に応じた効果的なデータ活用手法やリアルタイム解析技術は未整備であり、戦術的活用にも課題が残る。本研究では、選手のトラッキングデータを用い、攻撃がシュートに至るかを予測するモデルを構築する。このモデルにより、シュート成功に共通するパターンや重要なアクションを明らかにし、得点機会を増加させる戦術的指針を提供することを目指す。

2, 提案手法

本研究では、サッカー試合のトラッキングデータとイベントデータを用いて、攻撃の成否を予測し、有効なアクションを抽出する手法を提案する。過去試合のデータから選手位置やプレーシーケンスを抽出し、特徴量を加工後、ロジスティック回帰やRNNによる分類モデルを構築・評価する。最良モデルを用いてシュート確率を計算し、確率向上に寄与したアクションを「有効なアクション」として特定。これを基に攻撃成功要因を分析し、チームや選手への具体的な戦術改善案を提示する。最後に提案手法の有効性を定量評価し、攻撃成功率向上を検証する。

3, 実験

本章では、ロジスティック回帰とRNN(LSTM)を用いて、サッカーの攻撃がシュートに至る確率を予測するモデルを構築した。ロジスティック回帰にはトラッキングデータとイベントデータを用い、イベントデータからシュート(SHOT)とボールを失う(BALL OUT, BALL LOST)シーンを抽出。その瞬間および10秒前の選手全員とボールの座標を特徴量とした。一方、RNNでは、攻撃ごとにトラッキングデータを配列化し、0.04秒ごとの選手全員とボールの座標を特徴量として時系列データを使用した。データ前処理では、欠損値をゼロ補完し、時刻情報(Time [s])を基準にデータを昇順に並び替えた。さらに、HomeチームとAwayチームのデータの時刻をキーとして統合し、両チームの選手座標を同時に処理できる形式に整備。攻撃終了イベントは「BALL LOST」「BALL OUT」「SHOT」と定義し、それぞれの開始時刻を抽出した。攻撃区間は開始・終了時刻のペアに基づき分割し、(攻撃数, タイムステップ数, 特徴量数)の形状に整備。

攻撃数は332、タイムステップ数は最大4344、特徴量数は選手およびボール座標64とした。ラベル付けは、攻撃が「SHOT」で終了した場合を1(成功)、それ以外を0(失敗)として二値分類を行った。シュート確率の予測にはロジスティック回帰とRNN用い、過去のシュート成功・失敗のデータを用いてモデルを学習させる。その結果を基に有効アクションの抽出を行う。

4, 実験結果

機械学習を用いたシュート予測の精度比較表は表1のようになった。

表1: 精度評価比較

	クラス0 適合率	クラス1 適合率	クラス0 再現率	クラス1 再現率
①	0.93	0.02	0.87	0.33
②	0.94	0.00	1.00	0.00
③	0.97	0.31	0.55	0.92
④	0.81	0.17	0.45	0.50

- ① ロジスティック回帰による分析(アンダーサンプリングあり)
- ② ロジスティック回帰による分析(アンダーサンプリングなし)
- ③ RNNによる分析(アンダーサンプリングあり)
- ④ RNNによる分析(アンダーサンプリングなし)

5, 考察

評価に関して表1より、アンダーサンプリングありのRNNによる分析が最もバランスよく予測が出来ていると判断できる。

6, まとめ

本研究では、サッカーの攻撃成否を予測し、有効なアクションを抽出する手法を提案した。トラッキングデータとイベントデータを用いてロジスティック回帰とRNNを構築し、シュート確率の変化を基に最も効果的なアクションを特定した。結果として、縦パスやポジショニングの調整がシュート成功確率の向上に寄与することが明らかになった。さらに、抽出したアクションをアニメーション化することで、選手や指導者に具体的な指針を提供する可能性を示した。本手法は戦術の最適化やトレーニング支援に寄与する一方、データ偏りや守備要素の未考慮が課題である。今後は異なる試合データや守備要素の分析を通じ、汎用性と実用性の向上が期待される。