

ロジスティック回帰分析を用いた MMA の勝敗予測モデルの構築

Construction of an MMA Win-Loss Prediction Model using Logistic Regression Analysis

2241003 安部 拓人

Takuto ABE

指導教員 秋葉 知昭

This study predicts Mixed Martial Arts (MMA) about outcomes using past fight data from the Ultimate Fighting Championship (UFC). Bouts from four weight classes were analyzed, and a logistic regression model was constructed using relative differences in fighter statistics related to win rate, striking, and grappling performance. The results indicate that the difference-based model achieved the highest discriminative performance, suggesting the effectiveness of considering relative fighter abilities in MMA about outcome prediction.

1. 緒言

近年、スポーツ分野ではデータ分析や統計的手法を用いた試合結果予測に関する研究が進められている[1]。総合格闘技（Mixed Martial Arts : MMA）は、打撃技、組技、寝技といった複数の競技要素が同時に作用する競技であり、試合展開や勝敗が多様な要因によって決定される点に特徴がある。そのため、MMA の試合結果を事前に予測することは容易ではない。

本研究では、世界最高峰の MMA 団体である UFC を対象とし、試合前に取得可能な選手スタッツを用いた勝敗予測モデルを構築することで、MMA における勝敗要因を定量的に明らかにすることを目的とする。

2. MMA と UFC について

MMA は、打撃による KO や関節技・絞め技による一本勝ち、判定といった複数の勝敗決定方法を持つ競技である。また、体重別の階級制度が採用されており、階級ごとに試合展開や勝敗傾向が異なる。

本研究では、統一されたルール体系と豊富な公開データを有する Ultimate Fighting Championship (UFC) を分析対象とした。対象階級は、試合数が多く選手層の厚いフライ級、バンタム級、フェザー級、ライト級の 4 階級とした。

3. 使用データ

本研究では、UFC 公式サイトに掲載されている選手スタッツ[3]と、Wikipedia に掲載されている試合履歴データを基に分析を行った。対象選手は、各階級のランキング上位 15 名とし、2023 年までに実施された試合を学習データ、2024 年および 2025 年の試合をテストデータとして使用した。説明変数には、勝率、直近 3 試合の勝率、

KO/TKO 率、被 KO/TKO 率、有効打数（1 分あたり）、平均テイクダウン数（15 分あたり）などの指標を採用した。データ整理の結果、学習用データ 401 試合、テストデータ 105 試合を用いて分析を行った。

4. 分析方法

本研究では、勝敗を二値データとして扱い、ロジスティック回帰分析を用いて勝敗予測モデルを構築した[2]。MMA は対戦形式の競技であることから、選手の絶対的な能力よりも、対戦相手との相対的な能力差が勝敗に影響すると考えられる。

そこで、各説明変数について選手 A と選手 B の差分を算出し、これを説明変数として用いる差分モデルを主モデルとした。また、比較のために選手 A のスタッツのみを用いたモデルおよび勝率を除外したモデルも構築した。モデルの評価には、正解率に加えて ROC 曲線および AUC を用いた。

5. 分析結果

構築した 3 種類の勝敗予測モデルについて、テストデータを用いて性能評価を行った。はじめに、図 1 に各モデルの ROC 曲線を示す。ROC 曲線の比較から、選手 A のスタッツを用いた主モデルおよび選手間のスタッツ差分を用いた差分モデルはいずれも、ランダム分類を上回る識別性能を有していることが確認された。一方で、勝率を説明変数から除外したモデルでは、ROC 曲線が対角線付近に位置しており、勝敗を確率的に適切に識別できていないことが示された。特に、差分モデルは低い偽陽性率の領域において高い真陽性率を維持しており、対戦相手との相対的な能力差を考慮することが、勝敗予測において有効であることが示唆された。

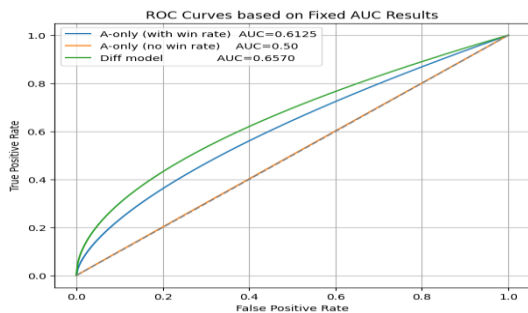


図1 ROC 曲線による比較

次に、図2に各モデルの評価指標を示す。正解率はいずれのモデルにおいても約0.8前後となり、大きな差は見られなかった。この結果から、正解率のみを指標として用いた場合、モデル構成の違いによる性能差を十分に把握できないことが示された。一方で、ROC 曲線に基づくAUCに着目すると、モデル間で明確な差が確認された。差分モデルは最も高いAUCを示し、次いで主モデル、勝率除外モデルの順となった。この結果は、MMAの勝敗予測において、予測確率の順位付け性能を評価する指標としてAUCが有効であることを示している。

TEST (2024-2025) accuracy: 0.8 AUC : 0.6125283446712018 test rows used: 105 / 105	DIFF MODEL accuracy: 0.8095238095238095 AUC : 0.6570294784580499	MODEL WITHOUT A_勝率 accuracy: 0.8 AUC : 0.5014172335600907 test rows used: 105 / 105
--	--	--

図2 3モデルの結果と比較

さらに、図3に差分モデルにおける各説明変数の回帰係数を示す。回帰係数の分析から、勝率差、有効打数差、テイクダウン数差といった指標が正の係数を示し、これらの変数が勝敗に対して正の影響を与える傾向が確認された。特に、勝率差は最も大きな係数を示しており、過去の試合結果を集約した指標がモデル内で重要な役割を果たしている可能性がある。一方で、被弾に関する指標や直近3試合の勝率差は負の係数を示しており、短期的な成績や被弾の多さが必ずしも勝利に直結しない場合があることが示唆された。これらの結果から、MMAの勝敗は単一の能力要因によって決定されるのではなく、打撃および組技に関する複数の競技特性の相対的な優劣によって左右されることが明らかとなった。

6. 結言

本研究では、総合格闘技団体UFCにおける過去の試合データを用い、ロジスティック回帰分析による勝敗予測モデルの構築を行った。対象階級を4階級に限定し、選手の勝率、打撃能力および組技能力を表す複数のスタツを説明変数として用いることで、MMAにおける勝敗要因を定量的に分析した。

分析の結果、正解率はいずれのモデルにおいても大きな差が見られなかった一方で、ROC 曲線およびAUCに基づく評価では、モデル構成によって識別性能に明確な差が生じることが確認された。特に、選手間のスタツ差分を用いた差分モデルは、他のモデルと比較して最も高いAUCを示し、試合結果を確率的に識別する能力が高いことが明らかとなった。この結果は、MMAの試合結果が選手の絶対的な能力のみで決定されるのではなく、対戦相手との相対的な能力差によって大きく左右される競技であるという特性と整合的である。

また、差分モデルにおける回帰係数の分析から、勝率差、有効打数差、テイクダウン数差といった指標が勝敗に正の影響を与える傾向が確認され、打撃および組技の双方が試合結果に関与していることが示唆された。一方で、被弾に関する指標や直近成績に関する変数が負の係数を示したことから、短期的な成績や被弾の多さが必ずしも勝利に直結しない場合があることも明らかとなった。

以上の結果から、本研究は、MMAの勝敗予測において選手の能力を個別に評価するだけでなく、対戦相手との相対的な差に着目したモデル設計が有効である可能性を示した点に意義があるといえる。今後の課題としては、対象データ数の拡充や階級ごとの詳細な分析に加え、試合内容の時系列データや他の機械学習手法を取り入れたモデルとの比較を行うことで、より高精度な勝敗予測モデルの構築が期待される。

参考文献

- [1] Sean R. Hackett, John D. Storey
Mixed Membership Martial Arts: Data-Driven Analysis of Winning Martial Arts Styles
- [2] Jiajie Yin
Fight Outcome Prediction Based on Fighter Statistics and Machine Learning Techniques
- [3] Ultimate Fighting Championship (UFC)
Official Fighter Statistics (公式サイト)