Leveraging Random Effects to Exploit Market Inefficiencies in UFC Betting

Jeremy Nichols, Arnav Mantro, Vinay Samuel

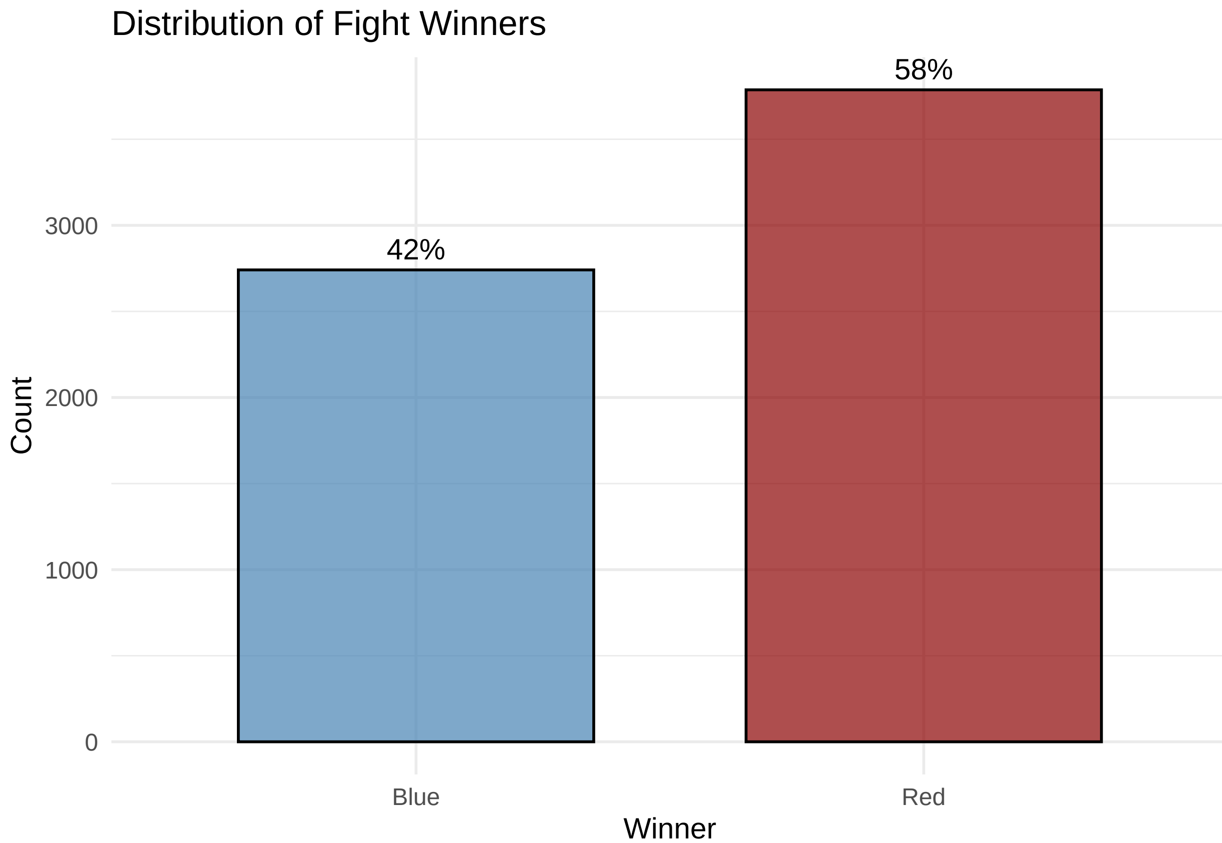
04/28/25

Introduction

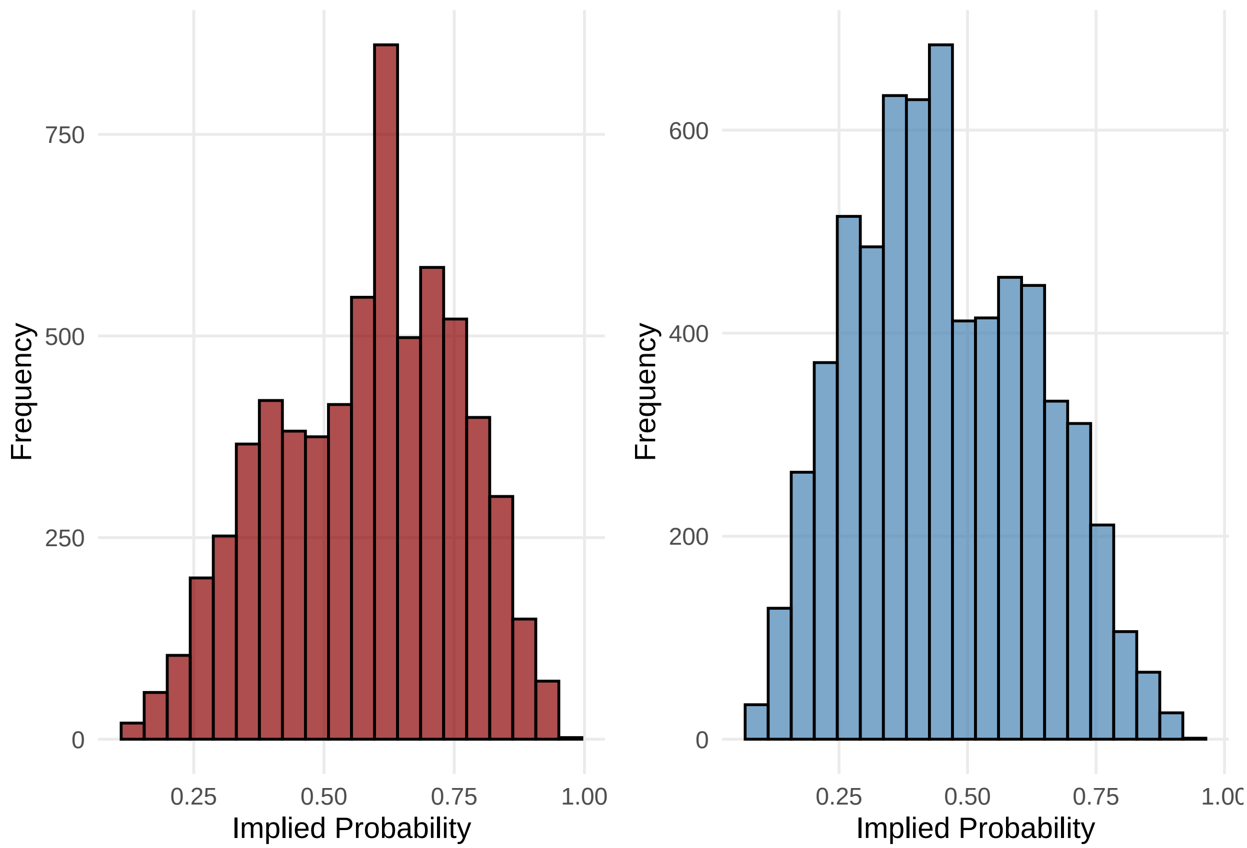
UFC bettors face an increasingly complex and volatile landscape, driving them to seek any analytical edge they can find when making picks. This challenge is especially compelling because the U.S. sports-betting market is booming— having roughly $149 billion in wagers during 2024—much of it online. Within that surge, UFC alone likely represents about $1-3 billion in bets, meaning even small improvements in prediction accuracy can translate into millions of dollars gained or lost. Understanding and modeling this problem therefore has huge financial stakes and offers a good testing ground for advanced analytics and machine-learning methods. It is inherently difficult to 'beat the market', as betting lines are set to reflect expected probabilities based on loads of data and include a spread for the books to make money. In our work, we will be exploring if multilevel logistic regression with random intercepts could enhance modeling accuracy and provide an edge against the markets.

Data

The dataset used for this project, called Ultimate UFC Dataset, was sourced from Kaggle. It contains statistics from UFC bouts, including rankings, fighter attributes, previous bout statistics, and betting lines. Some basic preprocessing was done in case of any missing values or incorrect dates, which consisted of imputing missing numeric values with medians, imputing categorical values with the mode, and converting date columns. Columns with lots of missing values or low relevance, such as "EmptyArena" and detailed fight finish descriptions were dropped. Rankings were also approached differently, where instead of having a dozen rank columns, which were mainly empty for most fighters, we created a variable to take the best rank out of the dozen possible columns for a particular fighter. As fighters typically fight in one weight class, their best rank should represent the highest rank in the weight class they are currently fighting at. Unranked fighters were given a default rank of 16, as ranked fighters will have values 0 (champion) to 15. First time UFC fighters, no matter the corner, were collapsed into a generic "OtherFighter" category. The exploratory data analysis (EDA) began by visualizing the distribution of fight winners, showing a slight bias towards red-corner fighters winning approximately 58% of matches compared to blue-corner fighters at just 42%.

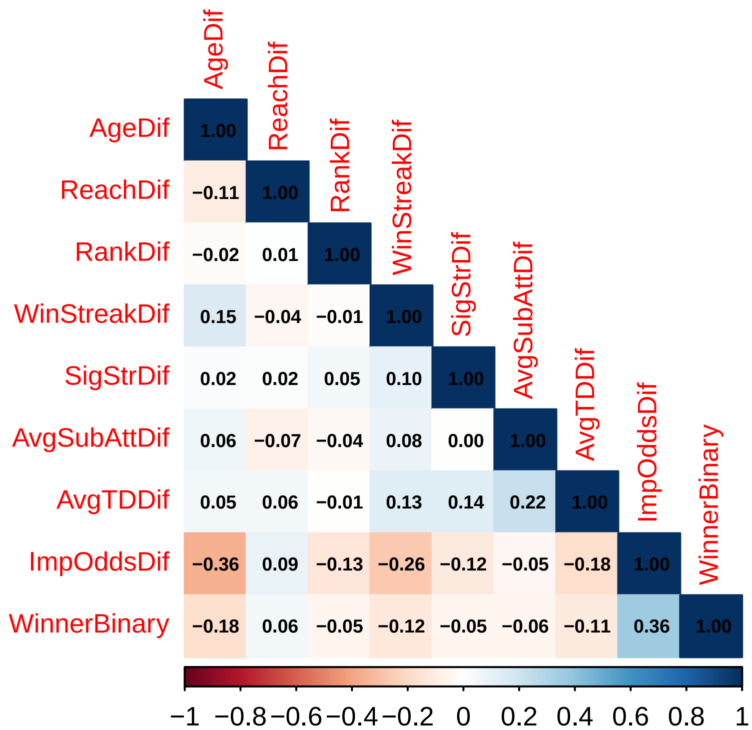


Further analysis involved examining the distribution of implied probabilities (based on betting lines) between for red and blue corner fighters. The left plot shows red corner fighters having implied probabilities of winning above 0.5, on average. The right plot for blue corner fighters shows the opposite, with the distribution shifter left of the 0.5 mark.



Lastly, we conducted a correlation analysis among key features representing differences between fighters, calculated as Red fighter minus Blue fighter. These included differences in best fighter rankings, win streak length, fighter age, average significant strikes per match, average submission attempts, average takedowns, as well as reach.

Correlation Matrix for Performance Features



We can see in the correlation plot above that the features that are higher correlated with red winning are age difference, win streak difference, takedown difference, and implied odds difference. For our predictors, we see age correlated with win streak, takedown difference correlated slightly with win streak, strike, and submission attempt differences. Moreover, implied odds differences appear to be correlated with age difference, win streak difference, and takedown difference more strongly. Given that our highest correlates have a value of 0.36, we dismiss any worries of collinearity between our predictors.

Methods

Model specification

We modeled the binary fight outcome (Red fighter win = 1, lose = 0) using a crossed random-effects logistic regression to account for repeated measures on individual fighters. Formally, for fight ij between Red fighter i and Blue fighter j:

$$
\begin{aligned}
\operatorname{logit}\big(P(Y_{ij}=1)\big) &= \beta_0 && (1)\\
&+ \beta_1\,\mathrm{RankDif}_{ij} + \beta_2\,\mathrm{WinStreakDif}_{ij} && (2)\\
&+ \beta_3\,\mathrm{SigStrDif}_{ij} + \beta_4\,\mathrm{AvgSubAttDif}_{ij} && (3)\\
&+ \beta_5\,\mathrm{AvgTDDif}_{ij} + \beta_6\,\mathrm{AgeDif}_{ij} && (4)\\
&+ \beta_7\,\mathrm{ReachDif}_{ij} + \beta_8\,\mathrm{ImpOddsDif}_{ij} + u_i + v_j, && (5)
\end{aligned}
$$

where

- $Y_{ij} \sim \text{Bernoulli}(p_{ij})$ with logistic link;
- $u_i \sim N(0, \sigma_u^2)$ and $v_j \sim N(0, \sigma_v^2)$ are fighter-level random intercepts;
- fixed effects β quantify the linear influence of standardized predictor differences on the log-odds.

Key assumptions include linearity of the log-odds in predictors, conditional independence of fights given random effects, and normally distributed random intercepts.

Alternative specifications and evaluation

We also fit two benchmark models:
1. **Simple logistic regression** with the same fixed effects but no random intercepts.
2. **Naïve market model** regressing outcome on implied-odds difference alone.

All models were trained on fights from January 1, 2016 to December 31, 2022 and tested on bouts from January 1, 2023 to December 31, 2024. Predictor differences were standardized (zero mean, unit variance) on the training split before application to the test set. The temporal split prevents information leakage and mimics real-world forecasting. We compare models by out of sample classification accuracy to assess reliability - metrics well-suited for binary sports-forecasting tasks.

Uncertainty quantification

We derived 90% confidence intervals for both our fixed and random effect coefficients via parametric bootstrap and nonparametric cluster-level bootstrap, respectively (1000 samples, each). This helped us balance computational feasibility with stable CI estimation in our crossed effects design, leading us to more conclusive results.

Results

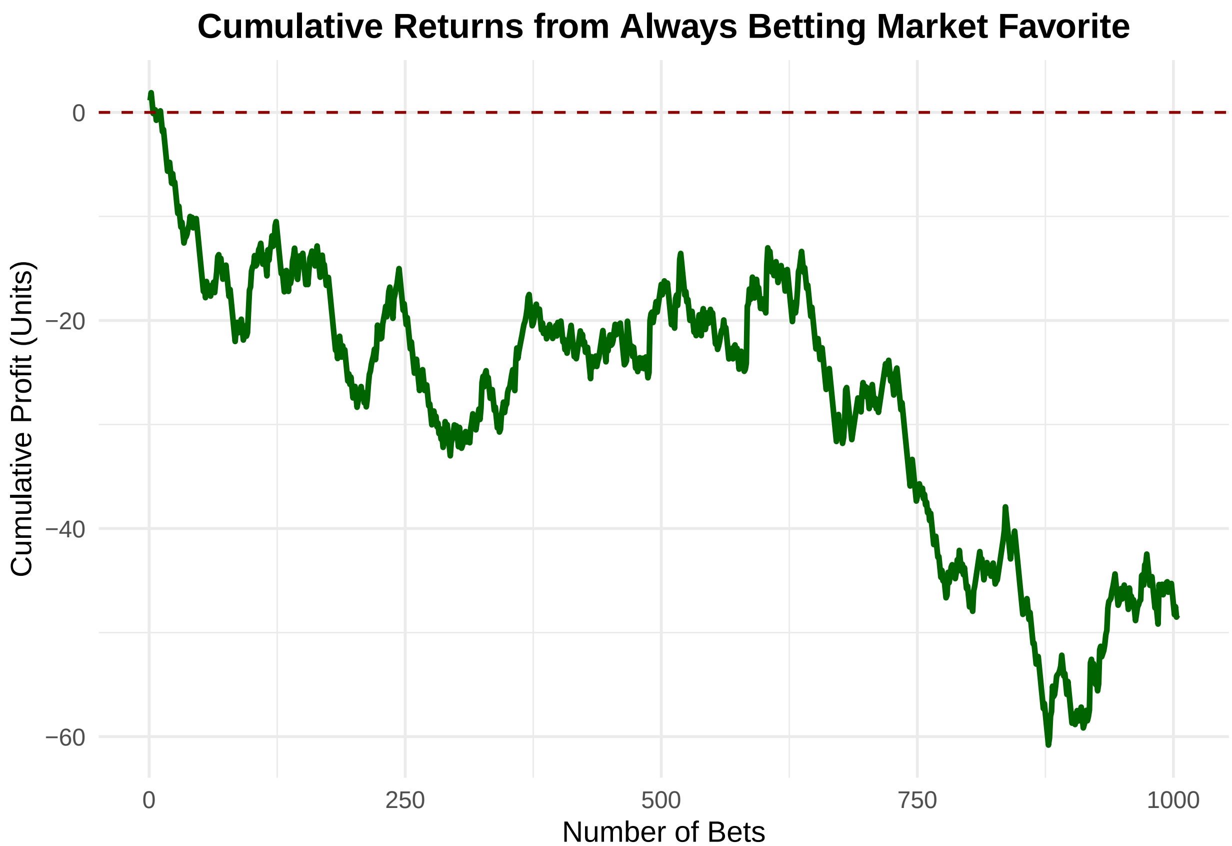
Our model output summary showed that the intercept, takedown difference, age difference, and implied odds difference to be statistically significant with $p < 0.01$. For each standard deviation increase in the red fighter's advantage in average takedown defense, the log-odds of the red fighter winning decrease, meaning that a higher takedown defense difference in favor of the red fighter is actually associated with a lower chance of winning. For each additional year the red fighter is older than the blue fighter, the log-odds of the red fighter winning decrease, so being older relative to the opponent reduces the red fighter's chance of winning. For each unit increase in the difference between the red and blue fighter's implied odds (in favor of red), the log-odds of the red fighter winning increase substantially, confirming that market odds are a very strong predictor of fight outcomes. The output residuals had mean very close to 0 and were symmetric, and correlation values between predictors were all below 0.3, supporting our assumptions.

Our crossed random-effects logistic model yields the highest out-of-sample accuracy on the 2023-2024 test split, correctly classifying 68.4% of fights, compared with 67.8% for a simple logistic regression and 66.7% for a naive market-favorite pick. Four our random effects model, we conducted one-season-out and one-weight-class-out cross validation. This returned nearly identical mean accuracy compared to our initial test accuracy.

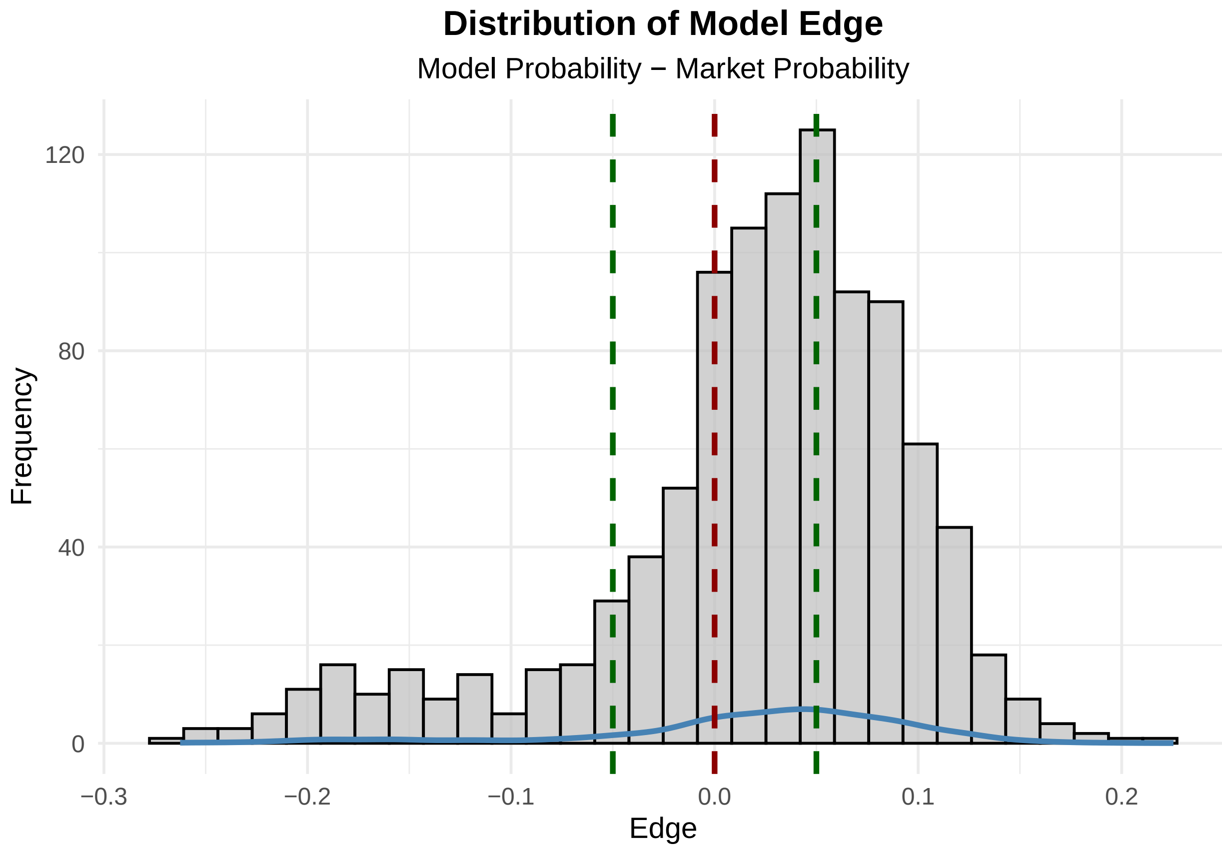
Table 1: Out-of-sample accuracy on temporal split

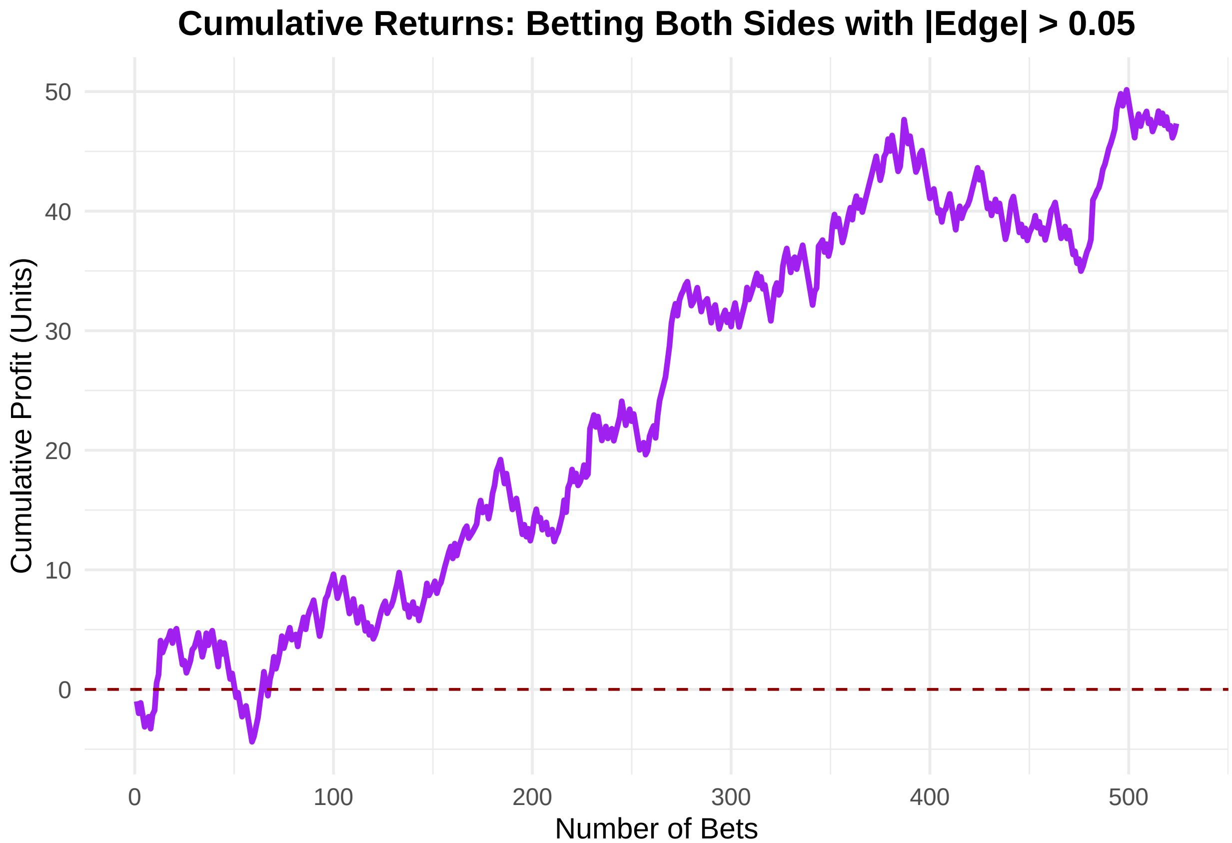
Model	Accuracy
Random Effects	0.684
Simple Logistic	0.678
Naive (Market Favorite)	0.667

Furthermore, under the table, we see the cumulative profit from always wagering on the market favorite over 1,000 bets (test set). The profit curve almost always consistently decreases, showing that betting for the favorite may not be the best choice.

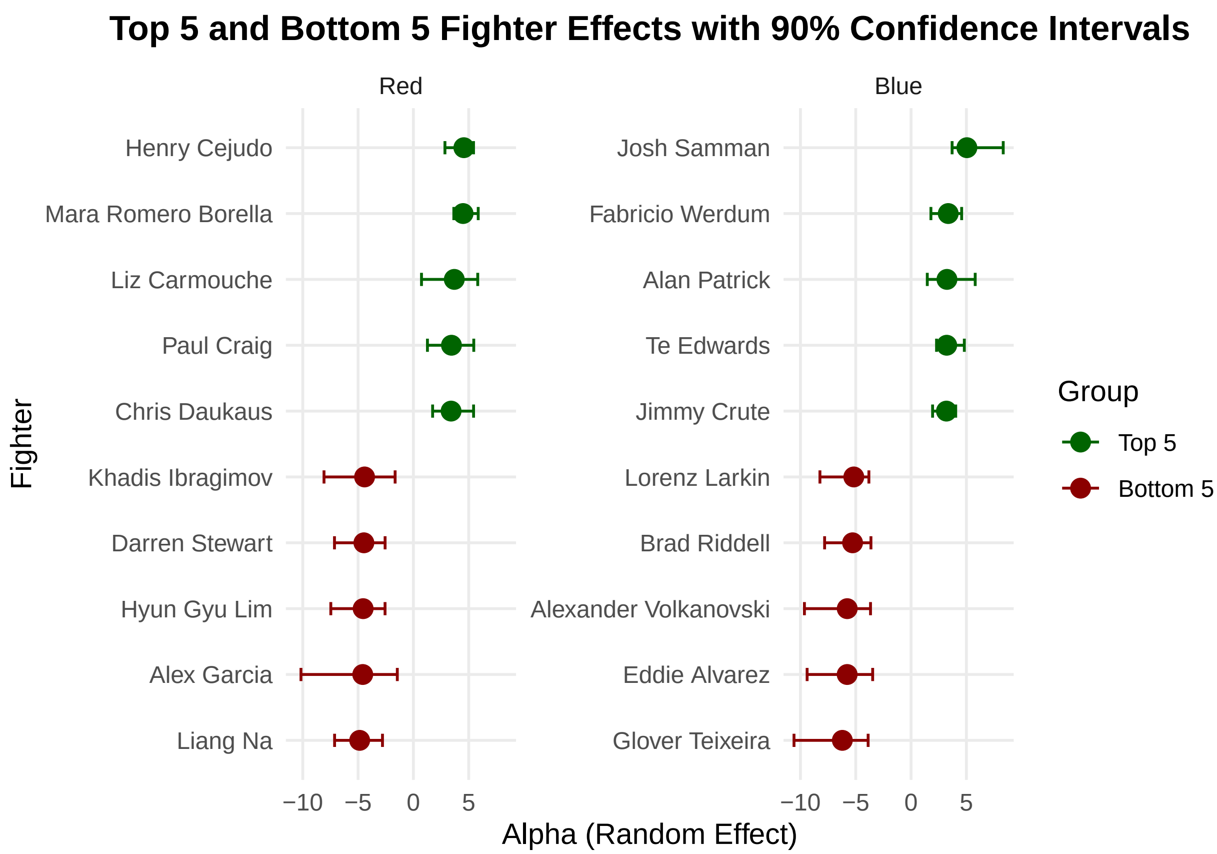


Turning to our model-driven strategy, we show the distribution of model edge (model probability - market probability). Most edges cluster within +/- 0.5, but a nontrivial tail of stronger negative and a lot of positive edges reveals where our model truly deviates from the market. Now, if we were to only bet on fighters with an absolute edge of over 0.05, we can expect significant cumulative returns with an ROI of about 10% per bet.


Distribution of Model Edge
Model Probability − Market Probability
Frequency
120
80
40
0
−0.3
−0.2
−0.1
0.0
0.1
0.2
Edge



Lastly, we plotted the top-5 and bottom-5 fighter random effects with their 90% confidence intervals. We see that there appears to be a significant difference in fighter random effects as top and bottom fighter intervals do not overlap. The top left group shows fighters who outperform the average when betting favorites, like UFC legend and former double-champ Henry Cejudo. The bottom right shows fighters who outperform the average as underdogs, like champions Alex Volkanovski and Glover Teixeira. These random effects for blue corner standouts are negative because our outcome variable is red corner winner. In the top right and bottom left corners we see fighters who perform worse than the average in their respective corner, with less recognizable names.



Discussion

Our crossed random-effects logistic model clearly outperforms both a simple logistic baseline and naïve market odds, highlighting the value of accounting for fighter-level differences in UFC fight outcomes. We must take into account the fact that UFC fighters usually fight 1-2 times per year, which means individual fighter data may be sparse and random effects may be heavily skewed. Other factors like undisclosed injuries, short-notice bouts, and potential differences between weight classes or genders are not included in our model and may limit our predictive power. To address these gaps, future work could incorporate rolling-window performance metrics, style and time-decay features, and compare male versus female divisions, while also exploring more advanced machine-learning architectures (e.g., neural networks, decision trees). Finally, responsibly testing our strategy in live betting markets would validate practical applicability and guide further model refinements. By integrating fighter-level random effects with pre-fight betting lines, we improve UFC outcome predictions - addressing our initial challenge of finding an edge in a potential $3 billion market.