



Mixed Membership Martial Arts: Data-Driven Analysis of Winning Martial Arts Styles

Other Sports
Paper ID: 1575

Sean R. Hackett and John D. Storey
Lewis-Sigler Institute for Integrative Genomics
Princeton University, Princeton, NJ 08544, USA

Abstract

A major analytics challenge in Mixed Martial Arts (MMA) is understanding the differences between fighters that are essential for both establishing matchups and facilitating fan understanding. Here, we model ~18,000 fighters as mixtures of 10 data-defined prototypical martial arts styles, each with characteristic ways of winning. By balancing fighter-level data with broader trends in MMA, fighter behavior can be predicted even for inexperienced fighters. Beyond providing an informative summary of a fighter's style, it is also the case that style is a major determinant of success in MMA. This is reflected by the fact that champions of the sport conform to a narrow subset of successful styles.

1. Introduction

Early events in Mixed Martial Arts (MMA) were touted as a chance to determine which styles of pure martial arts were most effective for defeating an opponent [1, 2]. These events, epitomized by the mid-1990s events of the Ultimate Fighting Championship (UFC) and Vale Tudo, featured matchups of stylistically diverse martial artists representing sports such as Boxing, Judo and Karate [1, 3]. As the sport of MMA has evolved, fighters have become increasingly well-rounded [1, 2]. Modern MMA fighters need to be effective strikers who dictate where the fight occurs, either by taking down their opponents or forcing them to fight standing.

Each MMA fighter's skills are drawn from a mix of multiple pure martial arts that define his/her personalized fighting style. While modern MMA fighters are more stylistically mixed than their pure martial artist forbearers, the mixtures of individual fighters differ. Some fighters would be referred to as strikers and others, as grapplers who specialize in martial arts such as Brazilian Jiu-Jitsu (BJJ) [4]. These stylistic characterizations are important when establishing match-ups; but to date, there is no quantitative procedure for characterizing the style of MMA fighters. Furthermore, because of the absence of such an approach, the impact of style on success in MMA has never been quantitatively investigated.

Every sport is enriched by the behaviors that distinguish its athletes and teams. Whether one is concerned with a pitcher's arsenal of pitches, the plays a football team employs or the locations from which a basketball player likes to shoot, style is an important, albeit often nebulous, concept in sports. To make style more accessible, data-driven modeling approaches will be invaluable. These models have enormous value – they can be used to identify prospects who are similar to superstars,

2016 Research Papers Competition
Presented by:



build a balanced team, and guide visualizations that inform casual fans of each athlete's strengths. To date, approaches that have aimed to define athletes' styles have generally attempted to identify latent factors that could serve as a representation of style [5, 6, 7]. These approaches have been restricted to spatial summaries of athletes (e.g., from where they score points). Consequently, their latent factors reflect common spatial patterns, like three-point shooting or dunking in basketball [6].

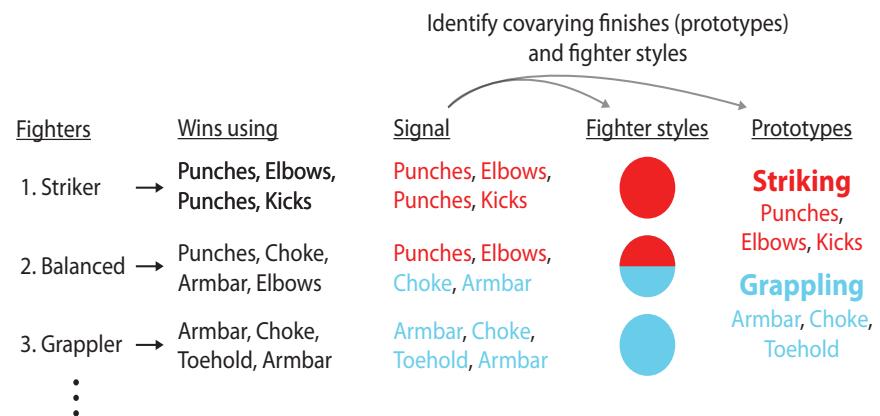
While latent variable based approaches have been successfully applied to sports where a large amount of athlete-specific performance data is available, such methods have the tendency to overfit patterns when little data is available or when complex patterns are sought (for example, if style were allowed to vary between seasons or between games). Most sports are not immune to such overfitting; for instance, the capacity for overfitting observed data has been well-demonstrated for the data-rich application of baseball batting averages [8, 9]. We would intuitively expect a batter with 300 hits out of 1000 at-bats to perform better in the future than a batter with 5 hits out of 10 at-bats, yet the second batter would have a higher batting average using the Maximum Likelihood Estimate (MLE) of batting averages. Indeed, it was demonstrated that the MLE was less accurate than the James-Stein estimate of batting averages, which shrinks observed batting averages towards the mean of all players based on the amount of individual-specific data [8]. Bayesian approaches of this sort have the capacity to balance athlete-specific data with overall behavior in order to improve prediction.

Characterizing the styles of MMA fighters is an unprecedented challenge because of both the amount and type of fighter-level data available. Most professional athletes compete in upwards of 100 games in a single season. In contrast, a successful MMA fighter may retire after only 20 bouts. With such limited data, a fighter's to-date behavior alone is insufficient to predict his/her future performance. Additionally, while the manner by which a fighter finishes a fight (e.g., finishes such as punches, kicks or chokes) is the purest depiction of his/her style [3], these metrics are categorical. Consequently, methods for defining spatially-oriented styles from other sports are not directly applicable. We address the above two challenges by applying latent Dirichlet allocation (LDA) [10, 11], a Bayesian mixed membership model that naturally accommodates categorical data and shares power across fighters, thereby balancing fighter-specific behavior with broader trends. The goal of this model is to estimate two mixtures that define a plausible underlying structure of fighters' styles (Figure 1): (i) Distinct martial arts (prototypes) have characteristic finishes and (ii) Fighters are a mixture of prototypes (with individualized weights summing to one).

Figure 1: Identifying fighter

styles based on

performance. We assume that an MMA fighter's wins generally reflect his/her training, either as a pure martial artist or as a practitioner of multiple prototypical styles. Each of these fighting styles utilizes a subset of possible finishes at characteristic frequencies.





Using data from >250,000 fights, we applied the mixed membership martial arts (3MA) model to identify 10 prototypical martial arts styles, each composed of a set of characteristic finishes. The style of each of ~18,000 fighters was defined based on his/her individualized combination of these prototypes. These stylistic makeups greatly improved the prediction of held out results. Beyond their value as an implicit readout of a fighter's abilities, fighter styles are a major determinant of success in the cage. In particular, winning fights in the upper echelons of MMA requires dynamic striking and the ability to "go the distance" by winning fights through judges' decision. Defending UFC champions, the most successful athletes in the sport, generally conform to these styles, providing insights into how fighting strategy affects success in MMA.

2. Data Set

Using the Sherdog Fight Finder [12], we obtained raw summaries of 257,582 previous amateur and professional bouts among 151,746 fighters through October 29, 2016. Each bout was summarized based on the pair of fighters, the result of the bout (win, loss, draw, no contest) and the way in which the bout was finished (e.g. by punches or unanimous decision). Rare finishes were manually combined with the most similar high-frequency finish to generate 50 distinct high-frequency categories (shown in Figure 2). Only fights that ended in a win for one fighter were considered, leaving 230,126 bouts for this analysis.

3. Identifying patterns in finish usage

At their most basic level, martial arts are sets of techniques that are used to win fights. Each fight is won by using a single technique, but if we look at a fighter's career, patterns may emerge where sets of finishes show up repeatedly across a fighter's victories. If these patterns are due to a martial art that is shared among fighters, we would expect sets of finishes to co-occur not only in a single fighter but also in a recurring manner across many individuals. For example, Muay Thai fighters use both elbows and knees, so we expect that fighters who win with elbows tend to win by knees. Thus, elbows and knees should co-occur.

To determine whether any subsets of the 50 finishes are frequently used in conjunction, we quantified how often each pair of finishes co-occurs (i.e. they are used to win fights by the same fighter), aggregating across all fighters. To answer this question, we compared observed co-occurrences of each finish pair (number of pairs of a finish x and y in a fighter's record, summed over fighters) with the expectation under independence. Then, using permutation testing, we determined how often such large deviations were expected. Out of 1,225 pairs of finishes, 289 pairs co-occur more frequently than expected, and 292 pairs co-occur less frequently at a false discovery rate of 0.05 [13].

Pairs of finishes that significantly co-occur are far from random; rather, they are organized into cliques of generally mutually correlated finishes (Figure 2). For example, kicks/knees form one tightly connected clique and leg submissions, another. From this analysis, we also see that patterns are more strongly influenced by the position from which a submission is applied, rather than by what type of submission is applied, *per se*. For example, the anaconda choke, applied from a front



headlock position, is independent from the triangle choke, which is applied from guard. In contrast, triangle chokes and armbars are frequently applied from guard and thus tend to co-occur.

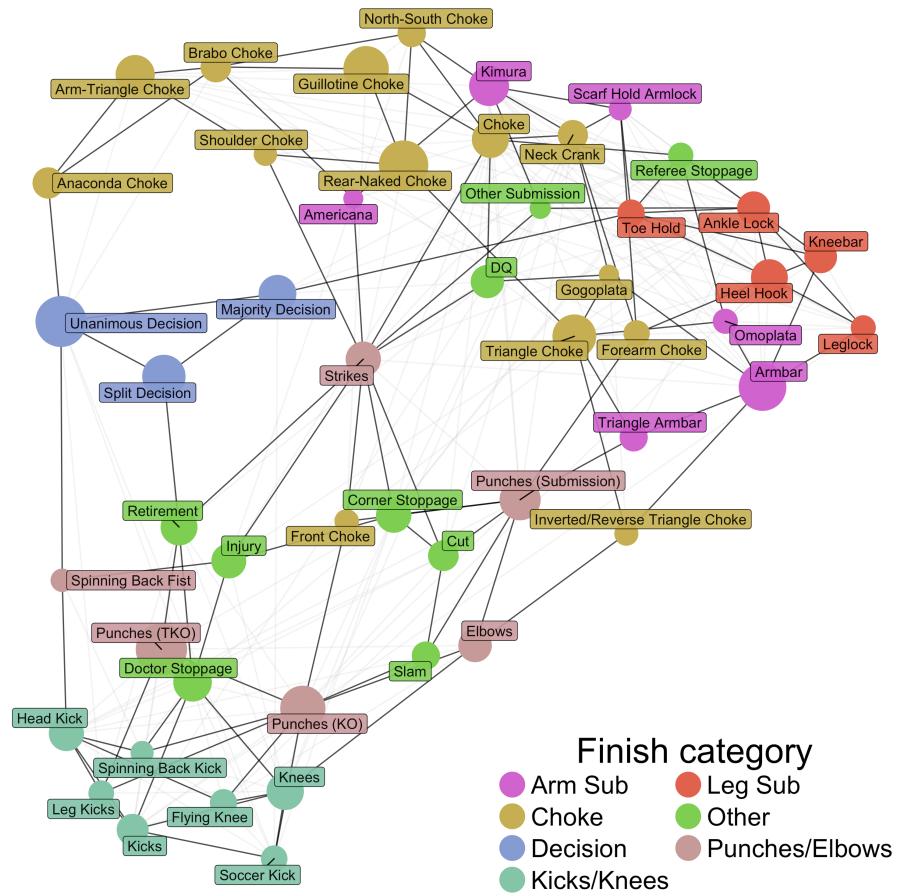


Figure 2: Structure of MMA finishes based on pairwise co-occurrence.
289 pairs of finishes significantly co-occurred among fighters. The three strongest connections to each finish (dark edges) were used to construct a layout; other significant edges are shown as faint background lines.

4. Determining fighter-specific styles

By aggregating over all fighters, we demonstrated that finishes exist in co-occurring groups. Directly applying this information at the fighter-level is challenging, however. Although we collectively have a large amount of information about fighters' performances, we have a more modest amount of fighter-specific data. No fighter has won a fight with each of the 50 possible finishes. In fact, the vast majority of experienced fighters do not even have 50 wins to date. Because we only have an incomplete readout of a fighter's capabilities, we need to balance what we know about a fighter with what we do not, hedging the uncertainty in fighter-specific data by using broader patterns across fighters. Specifically, we use the behavior of similar fighters to define common finish patterns and we then characterize individual fighters in the context of how frequently they have fought and the classes of finishes they have employed.

4.1. The mixed membership martial arts (3MA) model of fighter styles

In creating a model of fighter styles, our ultimate goal is to model the finish probabilities for each fighter. Statistically, this objective amounts to estimating fighter-specific finish probabilities: $\Pr(\text{Finish}_j | \text{Fighter}_i)$ for each finish $j = 1, 2, \dots, 50$ and each fighter $i = 1, 2, \dots, 17,778$ included in our



data set. (We restricted our analysis to fighters with four or more wins to date.) Our data is collected into an $I = 17,778$ by $J = 50$ matrix \mathbf{X} , where we note that the (i, j) entry is the total number of wins for fighter i of finish type j . We will denote $\omega_{ij} = \Pr(\text{Finish}_j | \text{Fighter}_i)$, which are the values we wish to estimate, and we collect these values into the $I \times J$ matrix Ω .

The most obvious estimate of the ω_{ij} values is the multinomial maximum likelihood estimate (MLE):

$$\hat{\omega}_{ij}^{\text{MLE}} = \frac{x_{ij}}{\sum_{j'=1}^J x_{ij'}}$$

However, when used on a per-fighter basis, the MLE is susceptible to overfitting the observed data (particularly for fighters with a smaller number of observations), and the MLE estimates that unobserved finishes have zero probability. It also does not “borrow strength” across the many fighters’ observed data.

Instead, in line with other latent variable based approaches (such as principal components analysis and factor analysis) [14], we assume a lower dimensional structure to Ω . By doing so, we are assuming that fighters’ individualized styles can be attributed to a set of prototypical winning styles which will be smaller than the number of $J = 50$ finish types. We will denote the number of prototypes by K , where $K < J$, and we will estimate the value of K from the data. We assume the following decomposition of Ω :

$$\begin{aligned} \Omega_{I \times J} &= \Theta_{I \times K} \Phi_{K \times J} \\ \omega_{ij} &= \sum_{k=1}^K \theta_{ik} \phi_{kj} \end{aligned} \tag{1}$$

where $\sum_{k=1}^K \theta_{ik} = 1$ for all i and where $\phi_{kj} \in [0, 1]$ for all k and j . The interpretation of this model is that the “prototypes” are captured by $\phi_{kj} = \Pr(\text{Finish}_j | \text{Prototype}_k)$. These prototypes are mapped to each individual fighter i through the values θ_{ik} where $\theta_{ik} = \Pr(\text{Prototype}_k | \text{Fighter}_i)$, where $\omega_{ij} = \sum_{k=1}^K \theta_{ik} \phi_{kj}$. For fighter i , the vector $(\theta_{i1}, \theta_{i2}, \dots, \theta_{iK})$ yields his/her mixture of prototypes, i.e., the fighter’s mixture of martial arts winning styles. This model decomposition is often referred to as a “mixed membership model” [10] in that each individual is composed of an individual-specific mixture of memberships to different classes. Here, the different classes are the prototypes.

This approach results in grouping of $J = 50$ finishes into a smaller number of prototypes K , and thus provides increased generalizability when predicting previously unobserved finishes. For example, submissions that target the legs (e.g., kneebar and ankle lock; Figure 2) should be associated with the same prototype. Because of this, a fighter with only kneebar submissions would have some weight in this “leg-submission” prototype. Therefore, we would predict that he/she also has the capacity to apply ankle locks. One limitation of this approach, however, is that if a fighter had never used leg submissions, he/she would have zero weight in the “leg-submission” prototype. As a result, we would predict that there is probability zero for him/her to perform any leg submission.

In order to allow for finishes contained in unrepresented prototypes and to also limit the influence of unrepresentative patterns of finishes, we can place a prior on mixture components, in line with



the earlier example of applying shrinkage to batting averages in baseball [8]. Latent Dirichlet allocation (LDA) is a popular Bayesian modeling approach for incorporating such a prior into the categorical matrix factorization shown in model (1), and furthermore, for fitting this model (1) to data through a technique called variational inference [10]. In LDA, Dirichlet priors with parameters $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_K)$ and $\beta = (\beta_1, \beta_2, \dots, \beta_J)$ are placed on Θ and Φ , respectively. LDA then utilizes variational inference to estimate the maximum *a posteriori* probability (MAP) of the posterior distributions of Θ and Φ , $\Pr(\Theta|\mathbf{X})$ and $\Pr(\Phi|\mathbf{X})$. The estimates $\hat{\theta}_{ik}$ and $\hat{\phi}_{kj}$ are set to be these MAP estimates, and we finally form $\hat{\omega}_{ij} = \sum_{k=1}^K \hat{\theta}_{ik} \hat{\phi}_{kj}$.

The values of β do not greatly affect inference on Φ because Φ is a relatively small matrix ($K \times J$) that is estimated by pooling data across all fighters [15]. In contrast, choice of α greatly impacts estimation of Θ [15]. Each component of α , denoted by α_k , can be interpreted as the number of effective observations (pseudocounts) [16] of a prototype k that supplements each fighter's history. The overall magnitude of α , which we denote by $\alpha^s = \sum \alpha_k$, reflects the degree to which fighters' prototypes will be determined by their own records (low α^s) versus the average of prototype frequencies across all fighters (high α^s). Treating α as asymmetric (meaning the α_k values across prototypes may differ) is another important consideration because this allows larger and smaller prototypes to exist rather than arbitrarily enforcing that prototypes be equally prevalent.

To approximate optimal values of α^s and K (since both the distribution of α and the values of β can be adaptively estimated), possible combinations of α^s and K were compared to determine which set generated an estimate of Ω that best predicted held out finishes. Quantitatively, for each pair of α^s and K values, we used 20-fold cross validation to estimate Ω for each of 20 subsets of the full dataset using the software MALLET [17]. We then calculated the log-likelihood of observations that were held out from each dataset:

$$\log(L) = \sum_{f=1}^{20} \sum_{i \in \mathcal{S}_f} \sum_{j \in \mathcal{S}_f} x_{ij} \log(\hat{\omega}_{ij}^s) \quad (2)$$

where \mathcal{S}_f are the indices of the test set for fold f and $\hat{\omega}_{ij}^s$ is the estimate based on the training set of fold f .

Parameter sets with high K generally resulted in held out $\log(L)$ similar in magnitude to those with moderate K , but high K models were less interpretable. Accordingly, we adopted a parsimonious approach to choosing a set of parameter values (small K , small α^s) among those with good performance ($\geq 95\%$ of the difference between $\max_{K, \alpha^s}(\log L)$ and multinomial likelihood model containing no fighter-specific information). This parsimonious approach resulted in setting $K = 10$ prototypes (two pairs of prototypes were combined because they contained the same majority finish) with $\alpha^s = 16$.



4.2. Performance and generality of 3MA: wins vs. losses and amateurs vs. the elite

To evaluate the performance of 3MA and assess its broader applicability, we applied the above approach both to elite subsets of fighters and to datasets constructed from fighters' losses rather than their wins (Table 1).

| Dataset | Type | Inclusion Criteria | N _{fighters} | N _{bouts} | K | α^s |
|-----------------|--------|---|-----------------------|--------------------|----|------------|
| All Data | | | | | | |
| all fighters | wins | 4+ wins | 17,778 | 142,775 | 10 | 16 |
| Subsets of Data | | | | | | |
| top fighters | wins | 4+ wins, 8+ fights, 1+ fights in "top promotions" | 2,973 | 40,244 | 11 | 14 |
| top vs. top | wins | 4+ wins, 8+ fights against top fighters | 408 | 3,673 | 14 | 22 |
| all fighters | losses | 4+ losses | 14,690 | 97,569 | 4 | 4 |
| top fighters | losses | 4+ losses, 8+ fights, 1+ fights in "top promotions" | 2,201 | 18,043 | 4 | 10 |
| top vs. top | losses | 4+ losses, 8+ fights against top fighters | 485 | 3,824 | 3 | 8 |

Table 1: 3MA was applied to six datasets constructed based on the fighters included and whether wins or losses were investigated. We considered "top promotion" to be UFC, Bellator, World Series of Fighting, PRIDE, DREAM and Strikeforce.

Our estimates of K and α^s are similar regardless of whether all fighters or only elite fighters are investigated, suggesting that regardless of the tier of competition, similar move sets are employed and accordingly amateur fighters may be important to better understand the elite. To quantitatively evaluate this assertion, for each of the six datasets described in Table 1, prediction accuracy was calculated for both inexperienced and experienced fighters (Figure 3).

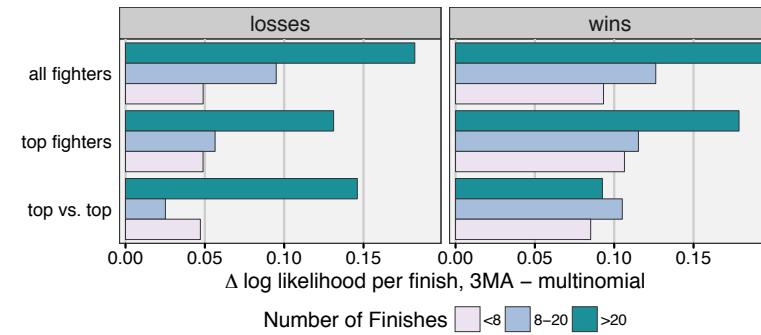


Figure 3: Performance of 3MA model relative to multinomial model where all fighters are assumed to use the same style (1 prototype).

In all cases, 3MA greatly outperforms a multinomial model, 1 prototype model (where finish probabilities for each fighter are proportional to finish frequencies across all fighters). In general, it is easier to predict how a fighter wins than how he/she loses (losing is not something over which a fighter has control). Additionally, the finishes of experienced fighters (those with >20 wins/losses) can be more accurately predicted when amateur fighters are included rather than using a smaller dataset composed exclusively of elite fighters. This suggests that patterns of MMA prototypes are nearly universal; accordingly, information from amateurs can improve the accuracy of defining the styles of champions. Because the complete dataset results in the best prediction of finish accuracy regardless of fighter experience, subsequent analyses will utilize all ~18,000 fighters.

We found $K = 3$ or $K = 4$ for losses compared to $K = 10$ for wins, indicating that fighters tend to lose in more generic ways than how they win. Losing prototypes are grouped into the large categories of



strikes, submissions, and decisions. Because a fighter's weaknesses can be faithfully read off from his/her record while winning prototypes are constructed from nuanced combinations of finishes, we focus our remaining attention on winning styles derived from analyzing all fighters. These styles are reflected in the Mallet software estimates of both the fighter-specific prototypes ($\hat{\Theta}$) and prototype-specific finishes ($\hat{\Phi}$).

4.3. Prototypes as mixtures of finishes

By studying the estimated values of $\phi_{kj} = \Pr(\text{Finish}_j | \text{Prototype}_k)$, we can characterize which finish types make up each prototype. We can also calculate the marginal probability $\Pr(\text{Prototype}_k) = \sum_{i=1}^I \hat{\theta}_{ik} / I$ as a summary of each prototype's overall relevance in the data. Summaries of $\Pr(\text{Prototype}_k)$ and $\Pr(\text{Finish}_j | \text{Prototype}_k)$ as estimated from the data can be found in Table 2.

| Prototype | Description | Pr(Prototype) | Finish Breakdown, Pr(Finish Prototype) |
|-----------|------------------------------------|---------------|--|
| P1 | Punches (TKO)/ Punches (KO) | 0.22 | 76% Punches (TKO), 20% Punches (KO), 3% Punches (Submission) |
| P2 | Unanimous Decision | 0.18 | 82% Unanimous Decision, 8% Split Decision, 7% Majority Decision |
| P3 | Rear-Naked Choke | 0.11 | 100% Rear-Naked Choke |
| P4 | Unanimous Decision/ Split Decision | 0.09 | 48% Unanimous Decision, 26% Split Decision, 16% Punches (TKO), 11% Other |
| P5 | Armbar/ Leg Sub | 0.09 | 72% Armbar, 11% Heel Hook, 5% Kneebar, 12% Other |
| P6 | Kicks/Knees/ Punches (TKO) | 0.08 | 27% Punches (TKO), 17% Punches (KO), 14% Knees, 42% Other |
| P7 | Triangle Choke/ Armbar | 0.07 | 56% Triangle Choke, 38% Armbar, 3% Rear-Naked Choke |
| P8 | Guillotine Choke/ Rear-Naked Choke | 0.07 | 84% Guillotine Choke, 15% Rear-Naked Choke |
| P9 | Rear-Naked Choke/ Kimura | 0.06 | 33% Rear-Naked Choke, 25% Kimura, 25% Arm-Triangle Choke, 17% Other |
| P10 | Punches (Submission)/ Choke | 0.05 | 35% Punches (Submission), 19% Choke, 13% Strikes, 33% Other |

Table 2: Summary of LDA-derived MMA prototypes. $\Pr(\text{Prototype})$ is given by the overall estimated frequency of each prototype in fighters ($\sum_{i=1}^I \hat{\theta}_{ik} / I$). $\Pr(\text{Finish} | \text{Prototype})$ is determined from the values in $\hat{\Phi}$. The top finishes for each prototype are shown.

The prototype characterization based on 3MA reproduces much of the structure found in Figure 2 but breaks this structure into discrete prototypes (Figure 4). These discrete prototypes appear to capture meaningful styles used by MMA fighters:

- The major striking prototypes are P1 and P6. P1 contains TKO and KO wins by punches, and thus represents boxing. P6 contains punches as well as elbows, kicks and other diverse strikes which are heavily featured in Muay Thai, Karate, Tae Kwon Do and kickboxing.
- Prototypes P2 and P4 contain wins by judges' decision. P2 primarily reflects unanimous decision, cases in which a fighter dominated his/her opponent for the length of the fight and all judges agreed that he/she won. P4 has some weight in unanimous decision but also strongly represents split decisions: fights in which judges disagreed about who won.

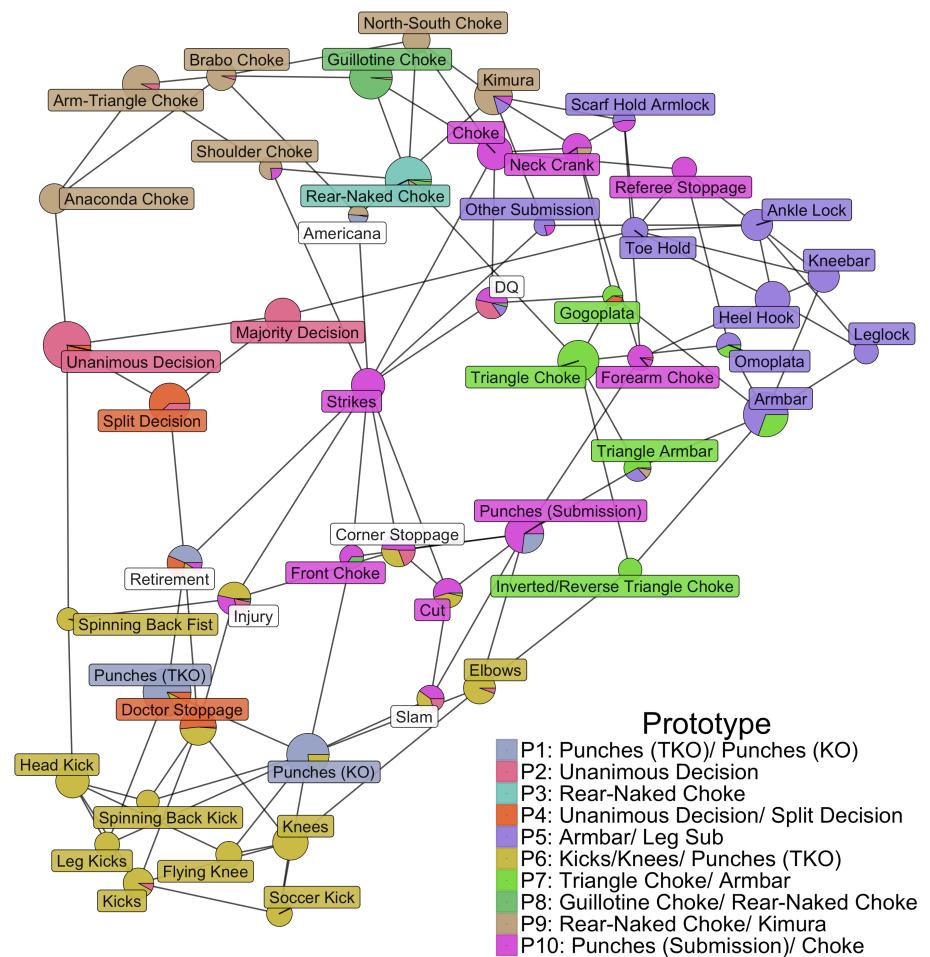


- BJJ and submission grappling are represented by five generally small probability prototypes. P5 primarily contains the armbar and submissions that target the legs. P7 are submissions from guard. P9 has some contribution from the rear naked choke but is primarily composed of submissions applied from side-control and front headlock. P8 and P3 each largely contain a single submission: the guillotine choke and rear naked choke, respectively. These submissions, while traditionally a part of BJJ, are known and practiced by most MMA fighters.
- The final and smallest prototype, P10, is a hodgepodge of unusual stoppages, generic terms and easily defended submissions that were relatively common in circa 2000 MMA but have decreased in frequency by over 10-fold since. Thus, P10 is a prototype which dates a fighter to the formative years of MMA.

This analysis does not include wrestling, judo and other martial arts that are specialized in achieving takedowns. Because these martial arts are not directly tied to MMA finishes, wins by their practitioners could manifest in many ways, such as through decision (P2) or submissions from strong control positions (P3, P8, P9).

Figure 4: 3MA identifies sets of finishes used together by fighters. Finishes from the network found in Figure 2 are colored according to $\Pr(\text{Prototype}|\text{Finish})$ where via Bayes' theorem we have:

$$\Pr(\text{Prototype}|\text{Finish}) = \Pr(\text{Prototype}) \times \Pr(\text{Finish}|\text{Prototype}).$$





4.4. Fighters as mixtures of prototypes

Prototypes exhibited a broad range of average frequencies, ranging from 5% to 22%. Individual fighters vary greatly about these averages, reflecting either their specialization or neglect of a given prototype and its associated finishes (Figure 5). Because all pairs of prototype loadings are negatively correlated, the mixtures of prototypes used by fighters are relatively independent; the weight in any one prototype does not necessarily entail higher weight in another prototype. While this trend would be partially expected due to tradeoffs between mixture components, the absence of strong associations between prototypes (as could be dealt with using correlated topic models [18]) suggests that the structure of finishes is appropriately discretized using LDA.

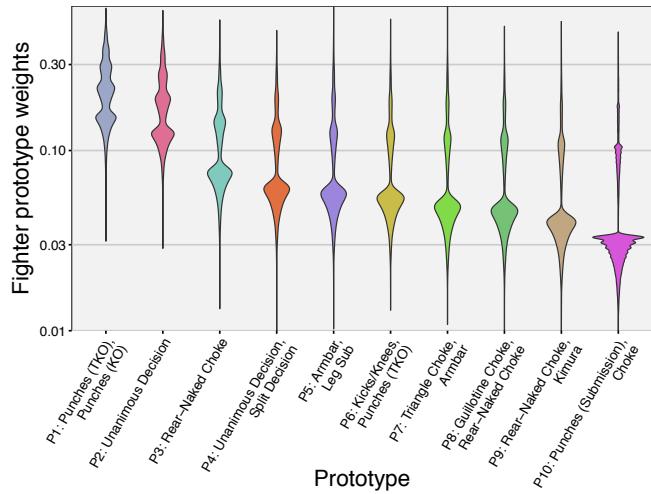


Figure 5: Violin plot showing the distributions of fighter prototype abundances .

5. Not all prototypes are created equal: winning styles in MMA

MMA fighters employ a variety of approaches to defeat their opponents, but it is unclear whether all styles are equally viable or if some have found disproportionate success in the cage. The success of a style may also be contextual. MMA fans frequently talk about an opponent being a “good matchup” for a fighter, reflecting that a fighter’s odds of winning might be influenced by both his/her own style as well as an opponent’s style. By creating a quantitative representation of style, we are able to systematically assess whether some styles are superior to others and whether this success depends upon an opponent’s style as well.

5.1. The role of matchups: prototypes form a strict dominance hierarchy

If some prototypes and their associated finishes are more important for winning MMA fights than others, we would expect that fighters who favor these finishes would win more frequently. The relative strength of a prototype may also depend on an opponent’s style. In such a case, the relative strength of a prototype may be contextual, resulting in intransitive properties like those seen in rock-paper-scissors.

To investigate how the prototypes perform against each other, we can consider who tends to win when all pairs of prototypes are matched against each other. Since every fighter is represented by a prototype mixture rather than by a pure prototype, we cannot directly count who wins during matchups. Instead, we consider the sets of bouts $\mathcal{B} = \{(b_1, b_2)\}$ where b_1 is the winning fighter and b_2 is the losing fighter (so $b_1, b_2 \in \{1, 2, \dots, I\}$), and characterize their prototype mixtures for each pair of prototypes m (in winning fighters) and n (in losing fighters) where $m, n \in \{1, 2, \dots, K\}$ and $m \neq n$. To the extent that $\hat{\theta}_{b_1 m}$ and $\hat{\theta}_{b_2 m}$ are large, prototype m has *de facto* defeated prototype n . More



generally, we weight each match based on $\hat{\theta}_{b_1m} \hat{\theta}_{b_2n}$, reflecting the extent to which prototype m has defeated prototype n . Then the score for winner m and loser n (s_{mn}) can be compared to the score for winner n and loser m (s_{nm}) to see if one prototype systematically dominates another:

$$s_{mn} = \sum_{(b_1, b_2) \in \mathcal{B}} \hat{\theta}_{b_1m} \hat{\theta}_{b_2n}$$

$$d_{mn} = \frac{s_{mn} - s_{nm}}{\frac{1}{2}(s_{mn} + s_{nm})} \quad (3)$$

To determine whether these dominance scores (d_{mn}) were significant, the observed value of each score was compared to 10,000 null values where winner and loser labels were permuted, and significant dominator-dominated prototype pairs were found at a false discovery rate of 0.05 [13].

Visualizing the dominances between all pairs of prototypes (Figure 6) suggests that intransitivities are a relatively minor phenomenon. Instead, prototypes occupy five tiers of dominance, wherein prototypes assigned to higher tiers dominate all prototypes in tiers beneath them. The top most successful tier of prototypes contains only P2 (unanimous decision), while the least successful tier includes P5 (armbar and leg submissions), P8 (guillotine choke and rear-naked choke), and P10 (punches submission and choke).

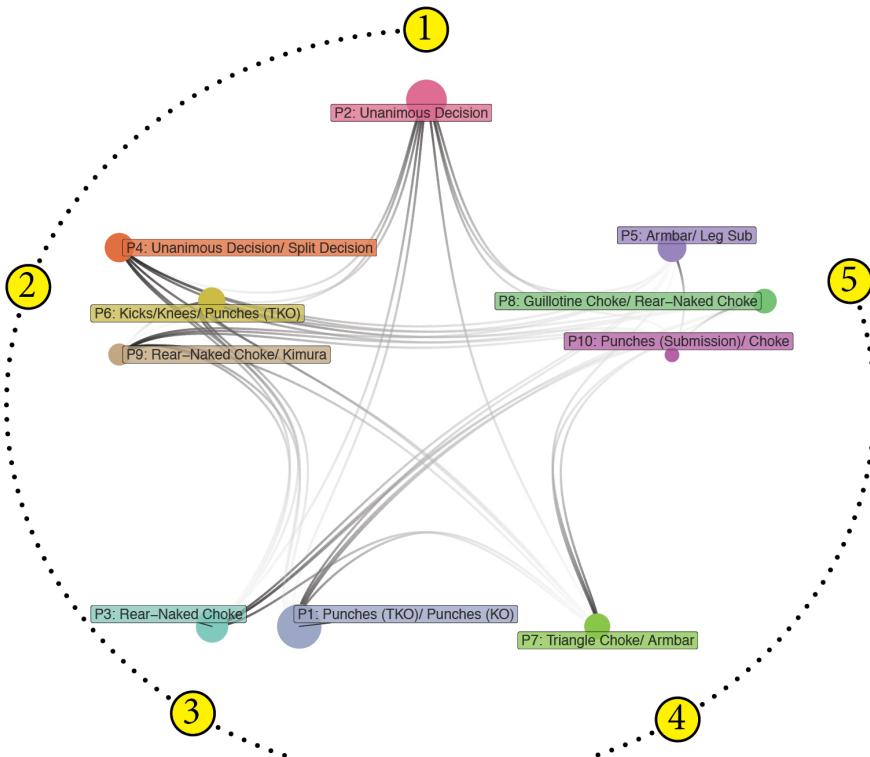


Figure 6: Prototypes form tiers of effectiveness.
 Network edges that indicate a significant dominance score link pairs of prototypes, pointing from a dominated prototype (light) to its dominator (dark). Groups of prototypes with mutually weak or insignificant dominance scores were grouped together into tiers, and tiers were arranged from the best performing (1) to the weakest performing prototypes (5).



5.2. Effect of prototype mixtures on overall success

Because prototypes are either effective or ineffective without regard to an opponent's strategy, each fighter's performance can be explained based on his/her prototype mixture, setting aside attributes of his/her opponent. To investigate prototype performance from the perspective of individual fighters, we sought to evaluate the degree to which fighter win percentages could be explained by prototype frequencies, i.e., $\text{Pr}(\text{win} | \hat{\Theta})$. One challenge with this strategy is that, by construction, fighters with extreme values of any prototype component are inherently veterans; they have won enough fights that their observed finishes have overwhelmed the influence of the prior. By contrast, fighters whose prototype proportions are near the prior proportions generally have very few fights. Therefore, their performance is more representative of an "average fighter." Because prototype values are partially conflated with experience, we need to tease out these two effects in order to determine the direct contribution of fighter styles on performance.

In order to separate the influence of experience and fighter styles ($\hat{\Theta}$), using generalized additive models (GAMs) with a logistic link function, the win percentage of each fighter i , $\text{Pr}(\text{win})_i$, was modeled based on the number of fights won (binned into intervals which include a similar number of fighters) and the 10 estimated prototype mixture components ($\log \hat{\theta}_i$). Logistic regression was used because $\text{Pr}(\text{win})$ can be treated as Binomial successes (wins) and failures (losses); accordingly, inference is best performed on the log odds of wins versus losses [19, 20]. The use of generalized additive models provides flexibility, allowing for the identification of nonlinear relationships between each explanatory variable and the log-odds of winning [21]. Fights were binned and mixture components were log-transformed in order to approximate Normal distributed values so that prediction was not unduly influenced by a minority of fighters.

From the fitted model (Figure 7A), we can see that the odds of winning increases monotonically with the number of wins as expected, while the influences of prototype abundances on the odds of winning, though highly significant ($p < 10^{-200}$, ANOVA versus a model excluding prototypes) are more complicated. In Figure 6, we noted that prototypes form tiers of dominance. Here, we see that a fighter's success with a prototype greatly depends on the degree to which the fighter has invested in it. Based on dominance, P2 is the most important prototype for success overall, but its benefits actually peak once a fighter contains 25% of this style and its effectiveness decreases at higher values. Aside from P2, at high values of other prototypes, the ranking of prototype effects on the odds of winning is similar to the tiers of dominance (Figure 6). Prototypes P5, P8 and P10 still negatively impact the odds of winning, while the remaining prototypes are all generally positive. These positive prototypes either have strong effects but impact few fighters (e.g., extreme values of P4, P6 and P9) or relatively weak effects that impact many fighters (e.g., intermediate values of P2 and high values of P1). To provide a sense of the magnitude of these effects, considering style alone, a 50%:50% P6:P9 fighter and a 50%:50% P5:P7 fighter would translate to an expected win percentage of 60% and 43%, respectively.

Experience and prototypes together strongly predict fighters' true win frequencies; but, the prediction is still quite remarkable with the role of experience removed (generating a model with experience and then correcting for this influence in the fit) (Figure 7B).

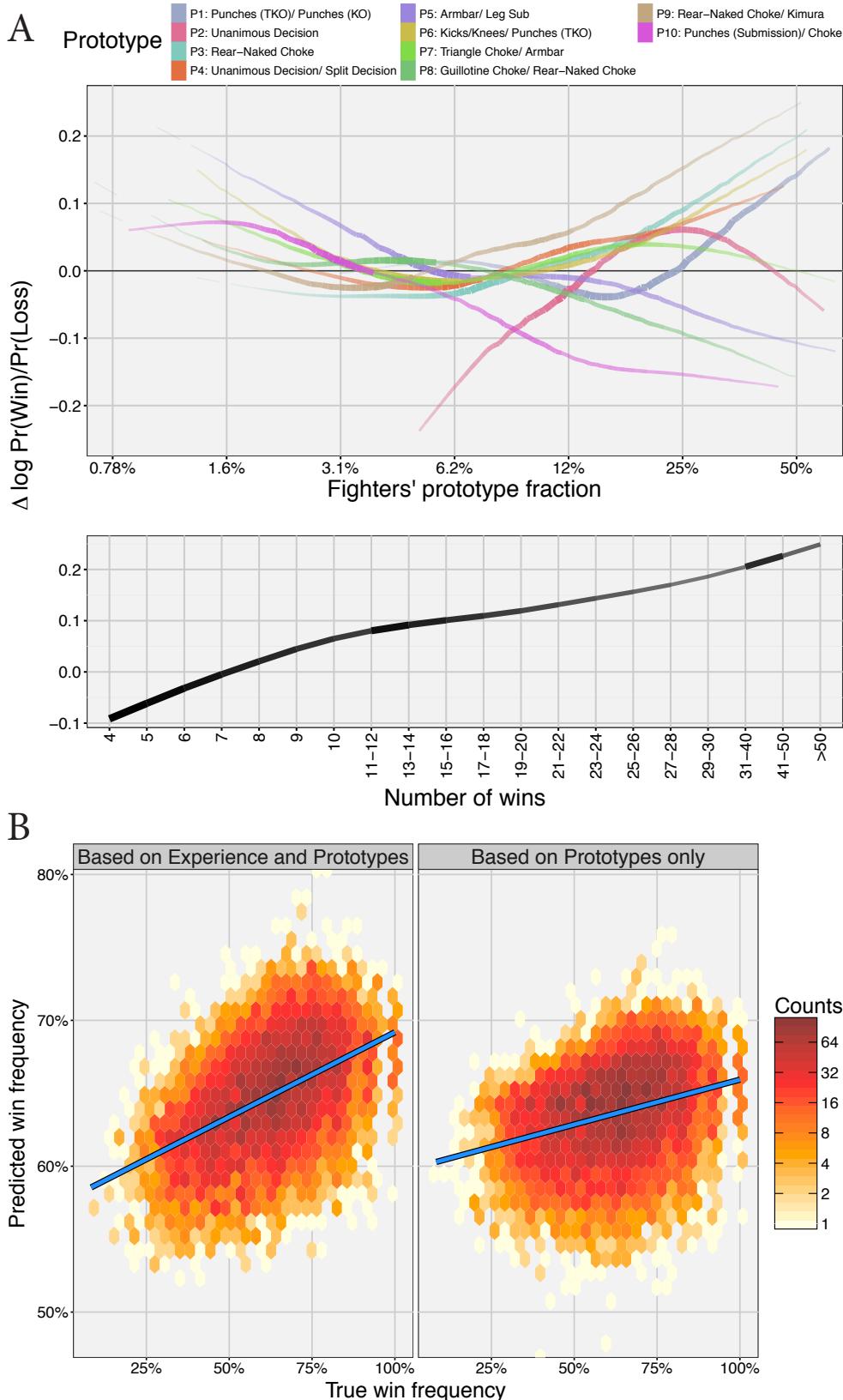


Figure 7: Predicting wins based on fighters' experience and prototypes. A)
The impact of prototype magnitude and number of fights on fighters' odds of winning (versus losing) was estimated using a generalized additive model. B) The fitted values of fighters' $\text{Pr}(\text{win})$ is compared to fighters' true win percentage. The impact of number of wins was subtracted from fighters' $\text{Pr}(\text{win})$ to generate an experience-corrected estimate based on prototypes only.



Because prototype specialization is an important predictor of how frequently each fighter wins, an interesting question is whether the most successful MMA fighters, UFC champions who have successfully defended their title at least once, utilize these effective prototypes. Looking at the predicted win frequencies of fighters (corrected for experience), defending UFC champions favor effective prototype mixtures (Figure 8A). Visualizing the most successful of these champions (those with the most title defenses), strong dynamic striking (represented by P6 and to a lesser extent P1) and the ability to “go the distance” by winning through decision (P2) are the major attributes that distinguish the best of the best (Figure 8B). These trends suggest that being a great fighter is not sufficient to become a champion; fighters must also embrace a style that potentiates their success.

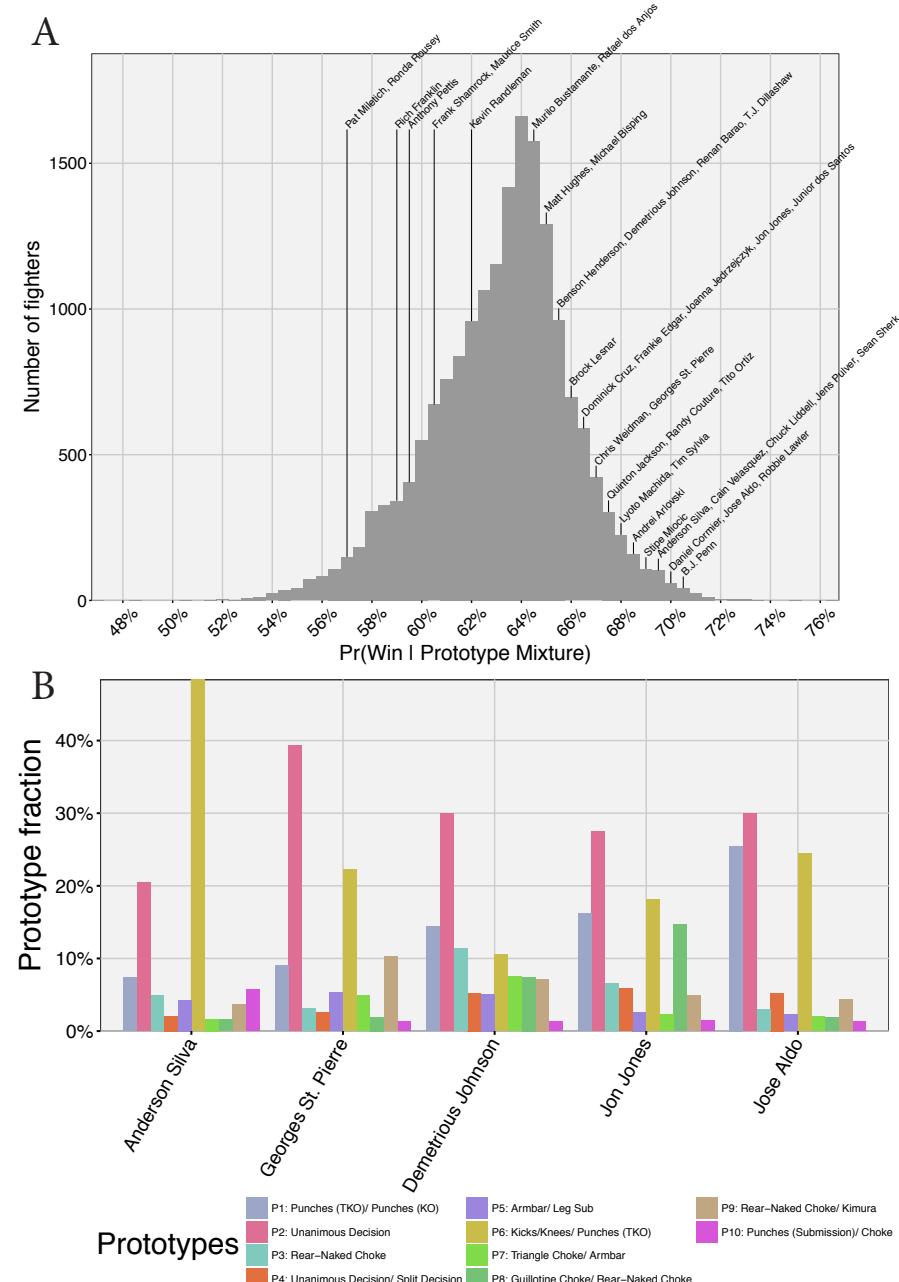


Figure 8: UFC champions favor the most effective prototypes. A) Distribution of fitted win probabilities for all fighters based on their estimated style, correcting for experience. Defending UFC champions (all listed names) heavily favor successful prototypes. **B)** Prototype mixtures of the 5 champions with the greatest number of title defenses.



5.3. Shedding light on success in MMA

Applying unsupervised approaches, we defined style in MMA and found that different elements of MMA fighters' games have realized varying success in the cage. Lessons emerge from this analysis.

Successful strategies: setting the pace and employing dynamic striking

- Being able to win by decision is a crucial component of every fighter's game. Winning by decision entails not only having sufficient cardio to fight for 15+ minutes but also doing so with enough of an advantage to be declared the victor. Thus, winning by decision may be an indicator of the overall fitness and caliber of a fighter. We also note that fighters who overspecialize in securing decision victories become less effective, thus securing decision victories should only be a part of a fighter's strategy rather than his/her primary goal.
- Winning by using kicks, elbows and knees is a greater indicator of success than victory by using punches, even though punches are the most frequent way of finishing a fight. This Muay Thai style could be advantageous for several reasons. Kicks, elbows and knees are important for controlling distance; they are generally acknowledged as being more powerful than punches, and they may also be an indicator of a top-level striker who would be dangerous with punches.
- The most successful submissions are the kimura, rear naked choke, and other chokes applied with the arms (excluding the guillotine). A common feature of these successful submissions is that they are applied from a position of very strong control (back, mount and side control). Because of this, if an opponent escapes from one of these submissions, the submitting fighter will still be left in a strong position.

Unsuccessful strategies: sacrificing position for submission

- Fighters who specialize in guillotine chokes or leg locks are less successful overall. A common feature of these classes of submissions is that should the application of these attacks fail, they generally leave the attacker in an inferior position.
- Guillotine chokes are generally applied by pulling guard (or being forced down during a takedown). As such, should the attack fail, an opponent is in a generally advantageous position to pass guard and secure a strong control position.
- Leg locks are also not applied from the traditional dominant grappling positions. Accordingly, when they fail, an opponent often has enough space to either secure a good position or to revert to a standing position.
- The most poisonous style was an ensemble of finishes that were common in MMA around the year 2000, but have since decreased greatly in prevalence. These finishes could be disadvantageous, but more likely, they simply date a fighter. If fighters have won by these finishes, it is likely that they are past the most competitive period of their careers and that this is reflected in their records.

While most modern UFC champions are strikers rather than grapplers and we point to subsets of the BJJ game that are ill-suited for MMA (the guillotine choke, leg locks, and to a lesser extent guard attacks), we find that fighters who specialize in submissions that are applied with strong positional control can be successful in MMA. Indeed, rather than committing to attacks that sacrifice position, the most successful BJJ practitioners in MMA (like Ronaldo "Jacare" Souza, Rafael Dos Anjos, and Demian Maia) blend BJJ with striking by challenging (or outright defeating) their opponents with



striking before moving to the ground. Then, when on the ground, these fighters apply chokes from strong positions.

While the manner by which a fighter wins is the purest reflection of his/her individualized style, and this summary can be obtained for a massive number of fighters, a fighter's style contributes not only to how he/she wins a bout but also to how he/she performs throughout the fight. Such data is increasingly captured, at least for top competitors, through analysis of video footage by organizations such as FightMetric [22]. These data could improve the accuracy with which style is defined for top fighters as well as capture other elements of a fighter's style, such as wrestling ability, that do not directly lead to characteristic finishes.

6. Broader applications in sports analytics

A common element of latent variable based approaches that have sought to define athletes' styles has been to find trends across athletes that summarize each athlete's performance. One limitation of such methods is that they are directly fit to the data, and thus do not account for stylistic components that have yet to manifest (or have, by chance, only been weakly demonstrated) in an athlete's performance. When applying principles in stylistic inference to MMA, a sport that is an extreme case of little athlete-specific data, it is crucial to share power across fighters using inference grounded in Bayesian modeling. While other sports are likely more insulated from such pathologies by virtue of possessing more per-athlete data, even when a large amount of data is available, directly fitting to per-athlete observed data might not most accurately estimate the true properties of the athlete (as has been demonstrated for baseball batting averages).

By harnessing latent variable based methods, we characterized MMA fighters' styles based on prototypes that naturally group finishes. Our analysis can be applied to any categorical measure of performance, including to other martial arts where fighters score points or win by using an expressive set of possible moves (e.g., BJJ, Tae Kwon Do, Sumo, Judo and Fencing). In each of these sports, the 3MA approach could help to systematize the styles of competitors and reveal factors governing success.

References

- [1] R. S. Garcia and D. Malcolm, "Decivilizing, civilizing or informalizing? The international development of Mixed Martial Arts," *International Review for the Sociology of Sport*, vol. 45, pp. 39–58, Mar. 2010.
- [2] S. H. Bishop, P. La Bounty, and M. Devlin, "Mixed Martial Arts: A Comprehensive Review," *Journal of Sport and Human Performance*, vol. 1, no. 1, pp. 28-42, 2013.
- [3] G. J. Buse, "No holds barred sport fighting: a 10 year review of mixed martial arts competition," *British Journal of Sports Medicine*, vol. 40, pp. 169–172, Feb. 2006.
- [4] A. Hirose and K. K.-h. Pih, "Men Who Strike and Men Who Submit: Hegemonic and Marginalized Masculinities in Mixed Martial Arts," *Men and Masculinities*, vol. 13, pp. 190–209, Sept. 2009.
- [5] K. Goldsberry and E. Weiss, "The Dwight effect: A new ensemble of interior defense analytics for the NBA," *MIT Sloan Sports Analytics Conference*, 2013.
- [6] A. Miller, L. Bornn, R. Adams, and K. Goldsberry, "Factorized Point Process Intensities: A Spatial Analysis of Professional Basketball," *Proceedings of the 31st International Conference on Machine Learning*, vol. 32, 2014.



- [7] X. Wei, P. Lucey, S. Morgan, M. Reid, and S. Sridharan, "The Thin Edge of the Wedge": Accurately Predicting Shot Outcomes in Tennis using Style and Context Priors," MIT Sloan Sports Analytics Conference, 2015.
- [8] B. Efron and C. Morris, "Data Analysis Using Stein's Estimator and its Generalizations," Journal of the American Statistical Association, vol. 70, no. 350, pp. 311–319, 1975.
- [9] S. T. Jensen, B. B. McShane, and A. J. Wyner, "Hierarchical Bayesian modeling of hitting performance in baseball," Bayesian Analysis, vol. 4, pp. 631-652, 2009.
- [10] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," the Journal of machine Learning research, vol. 3, pp. 993–1022, Mar. 2003.
- [11] D. M. Blei, "Build, Compute, Critique, Repeat: Data Analysis with Latent Variable Models," Annual Review of Statistics and its Applications, vol. 1, no. 1, pp. 203-232, Jan. 2014.
- [12] Sherdog, "Sherdog Fight Finder: <http://www.sherdog.com/stats/fightfinder>."
- [13] J. D. Storey and R. Tibshirani, "Statistical significance for genomewide studies," Proceedings of the National Academy of Sciences of the United States of America, vol. 100, pp. 9440–9445, Aug. 2003.
- [14] W. Hao, M. Song, and J. D. Storey, "Probabilistic models of genetic variation in structured populations applied to global human studies.,," Bioinformatics, vol. 32, pp. 713–721, Mar. 2016.
- [15] H. M. Wallach, D. M. Mimno, and A. McCallum, "Rethinking LDA: Why Priors Matter," Advances in Neural Information Processing Systems, vol. 23, pp. 1973–1981, 2009.
- [16] G. Heinrich, "Parameter estimation for text analysis," Tech. rep., University of Leipzig, 2009.
- [17] A. K. Mccallum, MALLET: A Machine Learning for Language Toolkit, 2002.
- [18] D. Blei and J. Lafferty, "A Correlated Topic Model of Science," The Annals of Applied Statistics, vol. 1, no. 1, pp. 17-35, 2007.
- [19] V. S. Y. Lo, J. Bacon-Shone, and K. Busche, "The Application of Ranking Probability Models to Racetrack Betting," Management Science, vol. 41, no. 6, pp. 1048-1059, June 1995.
- [20] S. M. Crowe and J. Middeldorp, "A Comparison of Leg Before Wicket Rates Between Australians and Their Visiting Teams for Test Cricket Series Played in Australia, 1977-94," The Statistician, vol. 45, no. 2, p. 255, 1996.
- [21] T. Hastie and R. Tibshirani, "Generalized Additive Models," Statistical Science, vol. 1, no. 3, pp. 297-310, 1986.
- [22] R. Genauer, "Fighting with Data: Creating Statistics from Scratch in Mixed Martial Arts," in MIT Sloan Sports Analytics Conference, 2012.