The Analysis and Classification of the Outcome of a UFC Bout

December 12, 2019

Multivariate Statistical Analysis Done By:

Joseph Gadbois
Vichetra Lim
Jacob Conlon

The Analysis and Classification of the Outcome of a UFC Bout

Joseph Gadbois | Vichetra Lim | Jacob Conlon

College of Natural Sciences and Mathematics
California State University - Long Beach
Long Beach, CA 90808

Abstract

The Ultimate Fighting Championship is an American company that promotes the sport of Mixed Martial Arts (MMA) as a form of entertainment, inspiration, and for the pure talent and drive these athletes contain. The main purpose of this study is to analyze the head to head career statistics and personal characteristics of professional fighters to find insights on what makes for a great athlete in the sport of MMA. This can also be looked at from the perspective of the athletes themselves, their coaches, and their teams to create a better training regime when training for upcoming fights. The analysis consists of three multivariate techniques including multivariate analysis of variance, principal component analysis, and discrimination analysis and classification. After rigorously analyzing the data, it revealed that the two variables that tend to result in winning a bout were takedown accuracy, which is the percent of takedowns attempted and were successful, and takedown defense. After training, the final model classified unseen data with 73.07% accuracy given equal prior probabilities and equal costs of misclassification.

Introduction

The Ultimate Fighting Championship, more commonly known as the UFC, is an American company that promotes all forms of Mixed Martial Arts (MMA). The UFC holds fights, or bouts, between professional MMA fighters for promotional and entertainment purposes. The roster of fighters for each event typically include the best professional fighters in the world, and there are many events where individuals need to pay to watch it live on television. With the abundance of fans, many of them will bet on the outcomes of the bouts and pay a lot of money in hope that their favorite fighters will win.

Data

The data are obtained from a website containing UFC data called FightMetric. The data began as two datasets, one containing fighter statistics 1,058 different fighters and the other containing information and results of 3,446 bouts from 2008 to 2018. The variables of interest from each dataset are listed in Table 1 and Table 2 below.

Table 1:	***Fighter's Data***	***Data Description***
Variable:	Data Type:	Description:
name	*Categorical*	*Name of the fighter*
X1	*Numeric*	*Height*
X2	*Numeric*	*Reach: length in inches of fingertip to fingertip at shoulder height*
X3	*Numeric*	*Significant Strikes Landed per Minute*
X4	*Numeric*	*Significant Striking Accuracy*
X5	*Numeric*	*Significant Strikes Absorbed per Minute*
X6	*Numeric*	*Significant Strike Defense*
X7	*Numeric*	*Average Number of Takedowns per 15 Minutes*
X8	*Numeric*	*Takedown Accuracy - % Attempted and Successful*
X9	*Numeric*	*Takedown Defense*
X10	*Numeric*	*Average Number of Submissions per 15 Minutes*
X11	*Numeric*	*Career Win Percentage*
X12	*Numeric*	*Age*

Table 2:	***Bouts Data***	***Data Description***
Variable:	Data Type:	Description:
fighter1	*Categorical*	*Name of fighter 1*
fighter2	*Categorical*	*Name of fighter 2*
Y	*Categorical*	*Winner of the bout, response variable, binary (win/lose)*
me	*Categorical*	*Method – 3 classes: Decision, KO/TKO, Submission*
wc	*Categorical*	*Weight Class of the fight*
year	*Categorical*	*Year of the fight*

Since the data were originally partitioned into two datasets, they were combined to create one full dataset. To avoid have two replicates of every variable from the fighter's dataset, the name of every fighter was matched up with the names of fighter 1 and fighter 2 in the bout's dataset. Equivocating the issue of having two replicates of every numeric variable, they were taken to be the difference of the value of fighter 1 and fighter 2, resulting in a positive value if fighter 1 had a higher value, negative value if fighter 2 had a higher value, and zero if the values are equal. Therefore, all numeric variables have uniform units of measurement being the difference between the value of fighter 1 and fighter 2.

Statistical Methodology

1. Multivariate Normal Distribution

When working with high dimensional data, it is often times very useful to approximate the population with the multivariate normal distribution and many multivariate statistical techniques are based on the assumption that the population has a multivariate normal density. For $p \geq 2$, the p-variate random vector is given by:

$$\boldsymbol{X}' = [X_1, \ldots, X_p] \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}), where\ \boldsymbol{\mu}\ is\ the\ mean\ vector\ and\ \boldsymbol{\Sigma}\ is\ the\ covariance\ matrix.$$

The random vector, $\boldsymbol{X}$ along with its mean vector $\boldsymbol{\mu}$ has the dimensions $px1$, and the corresponding covariance matrix, $\boldsymbol{\Sigma}$ is pxp. The p-dimensional normal density corresponding with $\boldsymbol{X}$ is given by:

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{p}{2}}|\Sigma|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}, -\infty < x_i < \infty$$

This density is a member of an exponential family of distributions and the point estimators of $\boldsymbol{\mu}\ and\ \boldsymbol{\Sigma}$ are given by $\overline{\boldsymbol{X}}\ and\ \boldsymbol{S}$ respectively. The sufficiency principle applies to these estimators since they come from an exponential family.

The assumption of multivariate normality is particularly useful data but needs to be verified before its properties are naively used. The central limit theorem is useful in this case, stating that for a large sample size of independent variables, then the sum of those variables can be concluded to be approximately normal. Since that is rather specific and cannot be applied to every situation, there are other methods of assessing the assumption of normality, including:

Observe Marginal Distributions of **X**:

- Examine a histogram looking for a symmetric bell-shaped curve for moderately large sample sizes
- Examine a dot plot for small sample sizes looking for the symmetric shape expected from univariate normal distribution
- Examine a QQ-Plot looking for the observations to lie along a generally straight line
- Perform a level α hypothesis test

Observe Bivariate Distributions:

- Examine a scatter plots for the points to form an elliptical shape expected from the bivariate normal density contours
- Plot squared distances against chi-square quantiles with p degrees of freedom looking for observations to lie generally along a straight line

2. Multivariate Analysis of Variance (MANOVA)

MANOVA is a method of comparing $g \geq 2$ population mean vectors at the same time to determine whether they differ or not. MANOVA assumed each population is independent and consists of a random sample of variables. Each independent population is assumed to be multivariate normally distributed. One-Way MANOVA, the populations are assumed to have common covariance matrices. The test for equality of means is given by:

$$H_0: \boldsymbol{\tau_1} = \boldsymbol{\tau_2} = \cdots = \boldsymbol{\tau_g} = 0\ \ vs.\ H_1: not\ all\ \boldsymbol{\tau_i} = \boldsymbol{0}$$

Model:

$$\boldsymbol{X_{lj}} = \boldsymbol{\mu} + \boldsymbol{\tau_l} + \boldsymbol{\epsilon_{lj}}, where\ \epsilon_{lj} \sim N_p(\boldsymbol{0}, \boldsymbol{\Sigma})\ and\ \sum n_l \boldsymbol{\tau_l} = \boldsymbol{0}$$

SSP:

$$\boldsymbol{B} = \sum n_l(\overline{\boldsymbol{x}}_l - \overline{\boldsymbol{x}})(\overline{\boldsymbol{x}}_l - \overline{\boldsymbol{x}})' \text{ is for treatments and is the between SSP}$$

$$\boldsymbol{W} = \sum\sum(\boldsymbol{x}_{lj} - \overline{\boldsymbol{x}}_l)(\boldsymbol{x}_{lj} - \overline{\boldsymbol{x}}_l)' \text{ is for residuals and is the within SSP}$$

Wilks' Lambda: $\Lambda = \frac{|\boldsymbol{W}|}{|\boldsymbol{B}+\boldsymbol{W}|}$

Reject H_0 if $-\left(n - 1 - \frac{(p+g)}{2}\right)\ln(\Lambda) > \chi^2_{p(g-1)}(\alpha)$

3. *Principal Component Analysis*

Principal Component Analysis is a statistical technique concerned with reducing the dimensionality of the data by explaining the variation and covariation in the data through a few independent linear combinations of the original variables called principal components.
Given a covariance or correlation matrix from a p-variate random vector **X**, find the corresponding eigenvalues λ_i and eigenvectors $\boldsymbol{e}_i$.
The principal components are given by:

$$Y_1 = \boldsymbol{e}_1\boldsymbol{X} \rightarrow Variance\ Explained\ by\ first\ PC \rightarrow \frac{\lambda_1}{\sum\lambda_i}$$

....

$$Y_p = \boldsymbol{e}_p\boldsymbol{X} \rightarrow Variance\ Explained\ by\ p^{th}\ PC \rightarrow \frac{\lambda_p}{\sum\lambda_i}$$

4. *Linear Discriminant Analysis and Classification*

The purpose of Linear Discriminant Analysis is to achieve maximum separation between the two populations, π_1 *and* π_2 to classify new observations into the population in which they belong.
Max Separation $= D^2 = (\overline{\boldsymbol{x}}_1 - \overline{\boldsymbol{x}}_2)'\boldsymbol{S}^{-1}_{pooled}(\overline{\boldsymbol{x}}_1 - \overline{\boldsymbol{x}}_2)$
Fisher's Discriminant Function is given by: $\widehat{w} = \hat{y} - \widehat{m} \geq 0$

$$where\ \hat{y} = (\overline{\boldsymbol{x}}_1 - \overline{\boldsymbol{x}}_2)'\boldsymbol{S}^{-1}_{pooled}\boldsymbol{x}_0 \text{ and } \widehat{m} = \frac{1}{2}(\overline{\boldsymbol{x}}_1 - \overline{\boldsymbol{x}}_2)'\boldsymbol{S}^{-1}_{pooled}(\overline{\boldsymbol{x}}_1 + \overline{\boldsymbol{x}}_2)$$

Allocate $\boldsymbol{x}_0$ *to* π_1 if $\rightarrow \widehat{w} = \hat{y}_0 - \widehat{m} \geq 0 \rightarrow \pi_2$ otherwise.
Evaluating Error Rates: $APER$ *or* $\hat{E}(AER)$
APER $= \frac{sum\ of\ off\ diagonal}{total\ observations}$ $\hat{E}(AER)$ is found via cross-validation (holdout procedures)

Preliminary Analysis

1. *Explore Summary Statistics*

Before considering potential multivariate statistical methods, the simple statistics, covariance matrices, and correlation matrices for each population should be considered. There are 12 continuous variables measuring the difference of the values of fighter 1 and fighter 2 consisting of the differences in their career statistics and their personal characteristics, such as their height or age. There are 2 categorical variables, 'me' is the method of how the bout ended with class levels being DEC for decision, SUB for submission, and KO/TKO for knock out or technical knockout. The other is 'wc' for weight class consisting of 8 different men's weight classes. Those will be explored more in depth later. The response variable is Y, the winner of the bout.

From the correlation matrices, it is observed that the variables are not very highly correlated. The only pair of variables that have a relatively moderate correlation coefficient is X1 and X2 having a correlation of 0.65 in population1 and that same pair has a correlation of 0.66 for population 2. Other than that, the absolute value of all other correlations falls below 0.5.

In the covariance matrices, the majority of the variables and pairs of variables have low variance and covariance. The variance of X8 and X9 are extremely large though. The variance of X8 is 720.09 and the variance of X9 is 784.88 in population 1, and 751.88 and 800.17 respectively in population

2. In population 1, X8 and X9 have a covariance of 212.79 and 176.00 in population 2. All other covariances are less than 75 in both populations. Note that in population 1, the total variance is 1,907.32, and the total variation is 1,946.07 in population 2. That being said, X8 and X9 account for 78.9% of the total variation in population 1, and 79.8% of the total variation in population 2. The covariance and correlation matrices for each population can be found in the appendex.

2. Assessing the Assumption of Multivariate Normality

To assess the assumption of multivariate normality with twelve variables, it is said that for practical work it is typically sufficient to consider the univariate and bivariate distributions. In figures 1 and 2 shown below, population 1 is represented by the points in blue for fighter 1 and population 2 is represented by the points in red for fighter 2. The main diagonal shows the marginal distribution of each of the variables. Figure 1 on the left has the marginal distributions for X1 – X6, along with the bivariate distributions for those pairs of variables. Figure 2 on the right has the same information, except for it has the variables X7 – X12.

Figure 1

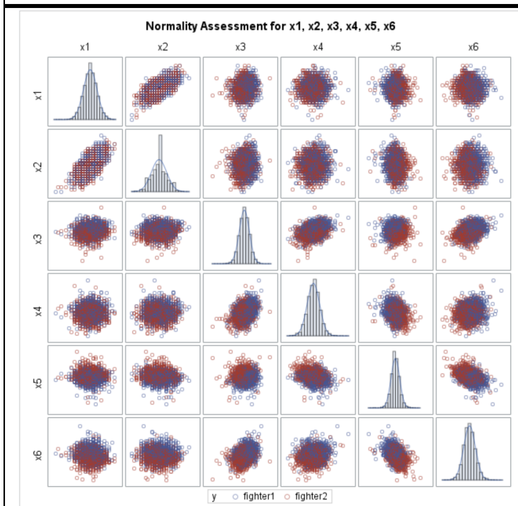
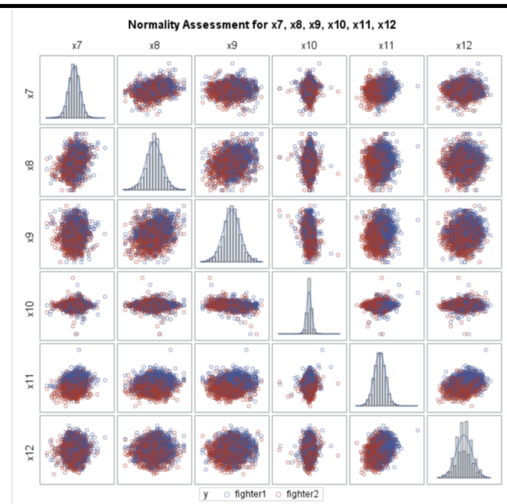


Figure 2



Considering the marginal univariate distributions, the histograms for all twelve variables seem to reveal a symmetric bell-shaped curve as expected from a normal distribution. The bivariate distributions look as if they represent an elliptical shape for both populations, indicating that the bivariate pairs shown above follows a bivariate normal distribution.

Since there are 12 variables and we observe two at a time for a bivariate distribution, that means there are 66 scatter plots we need to examine to see all of the pairs. A different method of determining the multivariate normality of these data is with a χ^2 plot.

Figure 3 for population 1 and figure 4 for population 2 below represent the χ^2 plots with $p = 12$ degrees of freedom. They are constructed by calculating the squared distances in ascending order and plotted against the χ^2 quantiles corresponding with 12 degrees of freedom. Notice how the points represent a fairly straight line that passes through the origin and has a slope of approximately 1. It may look like the slope is less than one but take note of the way the axes are spread out considering the domain and range and how much they differ. The line of points tends to be very close, if not right on the point that corresponds to $(X, Y) = (10, 10)$.

Figure 3

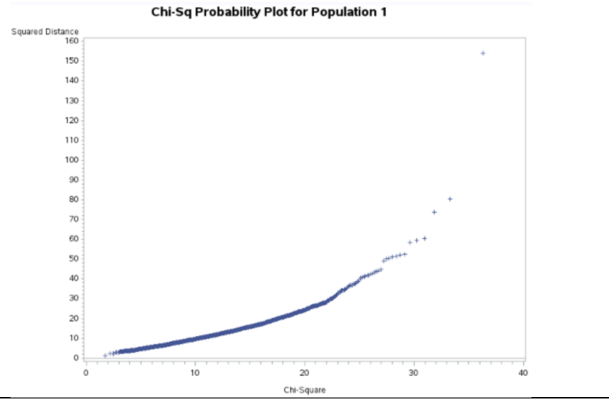
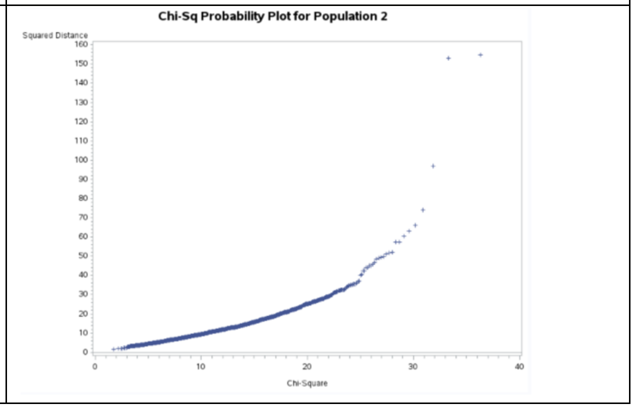


Figure 4



1. *Multivariate Analysis of Variance (MANOVA)*

As mentioned at the beginning, there are two categorical variables other than the response variable Y. There are a lot of factors that could play a role in predicting the outcome of a professional sports event. In the UFC, there are weight classes for a reason. If a bout was held where a professional fighter who weighs 170 pounds (welterweight) against a professional fighter who weighs 145 pounds (featherweight), that would be a 25-pound mismatch, and the likelihood of the featherweight fighter even phasing the welterweight would be minimal. Furthermore, one might expect that as the weight classes go up and down, the explosive power behind the punches, kicks, and takedowns as a whole would be maximized at a weight class where the fighters are heavy enough to throw very significant strikes, but agile enough to perform well on the ground with the wrestling, or grappling portion of the sport as well. The multivariate analysis of variance (MANOVA) determines whether or not there is a difference in the average value of each of the continuous variables at each level of the three categorical variables. The results from the analysis of Y=winner is shown in table 3 shown below:

Table 3:

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall y Effect
H = Type III SSCP Matrix for y
E = Error SSCP Matrix

S=1 M=5 N=1711

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.79250890	74.70	12	3424	<.0001
Pillai's Trace	0.20749110	74.70	12	3424	<.0001
Hotelling-Lawley Trace	0.26181548	74.70	12	3424	<.0001
Roy's Greatest Root	0.26181548	74.70	12	3424	<.0001

The hypothesis at the top of the table states that the mean value of Y=winner based on the 12 continuous X variables is the same for fighter 1 and fighter 2. This hypothesis is rejected because the probability of that happening, also known as the p-value, is very low. In fact, it is almost 0 as shown in the far-right column of the table where the header says $Pr > F$. There is clearly a difference in the averages between the winner and the loser of the fight.

The results of the analysis for me=method is shown in table 4 below:

Table 4:

MANOVA Test Criteria and F Approximations for the Hypothesis of No Overall me Effect
H = Type III SSCP Matrix for me
E = Error SSCP Matrix

S=2 M=4.5 N=1711

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.99204738	1.14	24	6848	0.2869
Pillai's Trace	0.00796478	1.14	24	6850	0.2872
Hotelling-Lawley Trace	0.00800410	1.14	24	5930.7	0.2867
Roy's Greatest Root	0.00593970	1.70	12	3425	0.0614

NOTE: F Statistic for Roy's Greatest Root is an upper bound.

NOTE: F Statistic for Wilks' Lambda is exact.

Unlike the results for Y, the MANOVA concludes that there is no statistically significant difference in the average values for the different methods of how the fight ends based on the 12 continuous X variables. A typical probability that the p-value in the far-right column needs to be lower than to reject the hypothesis of there being no difference in the averages is 0.05. All of the p-values in the far-right column are greater than 0.05, in this case the hypothesis is not rejected, and the conclusion is that there is no difference in means. Note that the p-value for the lowest row corresponding to Roy's Greatest Root is an upper bound, whereas the other three are lower bounds.

Finally, the results for weight classes are in table 5 shown below:

Table 5:

MANOVA Test Criteria and F Approximations for the Hypothesis of No Overall wc Effect
H = Type III SSCP Matrix for wc
E = Error SSCP Matrix

S=7 M=2 N=1711

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.97713089	0.95	84	20979	0.6217
Pillai's Trace	0.02306486	0.94	84	24010	0.6223
Hotelling-Lawley Trace	0.02320483	0.95	84	14696	0.6209
Roy's Greatest Root	0.00965980	2.76	12	3430	0.0010

NOTE: F Statistic for Roy's Greatest Root is an upper bound.

Since Roy's p-value is for an upper bound instead of a lower like the others, it means the opposite as the other p-values. The results from this analysis are the same as those for method, and the conclusion is that there is no significant difference in the averages of weight classes.

1. *Principal Component Analysis (PCA)*

Principal Component Analysis will be used to reduce the dimensionality of the data without losing too much information regarding the variance and covariance by creating a linear combination of the original variables that will contain most of the variability in just a few PCs.

Principal Components (PCs):

$$Y_1 = -.001X_1 + .002X_2 + .01X_3 + .08X_4 - .01X_5 + .11X_6 + .014X_7 + .63X_8 + .76X_9 - .006X_{10} + .073X_{11} + .02X_{12}$$
$$Y_2 = .004X_1 + .01X_2 - .01X_3 + .05X_4 - .01X_5 - .03X_6 + .02X_7 + .77X_8 - .63X_9 + .008X_{10} - .001X_{11} - .01X_{12}$$
$$Y_3 = .01X_1 + .02X_2 + .03X_3 + .25X_4 - .03X_5 + .24X_6 + .02X_7 - .08X_8 - .09X_9 + .01X_{10} + .92X_{11} + .105X_{12}$$

Variance Explained	Total Variation
(see table below)	Total Variance: 2008.9034936

Eigenvalues of the Covariance Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	1031.24784	459.19290	0.5133	0.5133
2	572.05494	403.22417	0.2848	0.7981
3	168.83077	68.43688	0.0840	0.8821

Total Variance	2008.9034936

Interpretation of Principal Components:

- Y_1: Can be thought of as a combination of X8 and X9 since it has a lot of weight for those two variables and essentially a negligible amount for the rest
- Y_2: Can be thought of as a combination of X8 and X9, it is very similar to Y_1 in how the absolute value of the weight are distributed through the variables
- Y_3: Can be thought of mostly as X11, with a little bit of weight for X4 and X6 as well

Recall the conversation about the covariance matrix from the beginning. The variables X8 and X9 account for around 79% of the total variation in the data. PC1 and PC2 account for 79.81% of the total variation in the data. That could potentially mean that the variation in the original data is explained, for the most part, by X8 and X9.

2. Discrimination and Classification

Discriminant analysis will return the final results of the analysis using the first three principal components from PCA. The procedure will test for equality of covariance matrices to find that the covariance matrices between the two populations are equal. Thus, a method called Fisher's Linear Discriminant Analysis will be used to create a rule to separate the data into their respective populations by trying to achieve the maximum distance between the two populations with the highest accuracy and the lowest error rates. Note that the cost of misclassification and the prior probabilities are equal.

The squared distance and the generalized squared distance in this case are the same, so table 6 shown below will be the generalized squared distance.

Table 6

Generalized Squared Distance to y		
From y	fighter1	fighter2
fighter1	0	0.83469
fighter2	0.83469	0

The classification rule for allocating a new observation into population 1 or population 2 is the function:

Fisher's Linear Discriminant Function:

Allocate $\boldsymbol{x}_0$ to population 1 if:

$$0.0195 * Prin1 - 0.00706 * Prin2 + 0.06197 * Prin3 - 0.01556 \geq 0$$

Allocate to population 2 otherwise.

We can find the APER, $\hat{E}(AER)$, and the Test Data Accuracy as shown below:

APER

Number of Observations and Percent Classified into y			
From y	fighter1	fighter2	Total
fighter1	1049 66.82	521 33.18	1570 100.00
fighter2	513 33.03	1040 66.97	1553 100.00
Total	1562 50.02	1561 49.98	3123 100.00
Priors	0.5	0.5	

$$\frac{521 + 513}{3123} = 33.11\%$$

$\hat{E}(AER)$

Number of Observations and Percent Classified into y			
From y	fighter1	fighter2	Total
fighter1	1048 66.75	522 33.25	1570 100.00
fighter2	513 33.03	1040 66.97	1553 100.00
Total	1561 49.98	1562 50.02	3123 100.00
Priors	0.5	0.5	

$$\frac{513 + 522}{3123} = 33.14\%$$

Test Data Accuracy

Number of Observations and Percent Classified into y			
From y	fighter1	fighter2	Total
fighter1	106 65.84	55 34.16	161 100.00
fighter2	32 19.75	130 80.25	162 100.00
Total	138 42.72	185 57.28	323 100.00
Priors	0.5	0.5	

$$\frac{106 + 130}{323} = 73.07\%$$

The apparent error rate and the expected actual error rate indicate an accuracy of around 66-67% but, when classifying data that was unseen, the model correctly classified observations 77.07% of the time.

Conclusion

After applying multiple multivariate statistical techniques, it is clear that the variables X8 and X9, which correspond to takedown accuracy and takedown defense play the biggest role in a UFC fighters' success. Even after principal component analysis, the first two principal components consisted of mainly the variables X8 and X9 and accounted for almost 80% of the total variation in the data. The principal component scores were then used as input in the linear discriminant function to ultimately classify unseen data with 73.07% accuracy. When working with data, there will always be error, and especially with data that has to do with sports it becomes even more unpredictable because there are so many unpredictable factors that can have effects on any given day.

Appendix

For Population 1

Covariance Matrix, DF = 1730												
	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12
x1	6.2598919	5.4574425	0.1901680	0.2071895	-0.0610526	-2.0433880	-0.4276836	2.0584797	0.4748319	0.2715656	0.8262617	2.0976568
x2	5.4574425	11.1232473	0.0415446	-0.2110902	-0.4023163	-2.8300321	-0.4078297	3.3920107	-2.7854530	0.3754098	2.1428085	2.6449184
x3	0.1901680	0.0415446	1.8997156	4.9205546	0.3206551	4.4269019	-0.4025460	4.2083562	10.2850072	-0.3127128	1.8419917	1.1598847
x4	0.2071895	-0.2110902	4.9205546	109.4039618	-2.8021061	8.8812341	2.2263271	59.1956632	27.2410334	0.4022398	15.1209416	-1.0983981
x5	-0.0610526	-0.4023163	0.3206551	-2.8021061	1.7885531	-5.1594178	-0.5753427	-6.3866769	-1.8014211	-0.1024584	-3.0137733	0.4922731
x6	-2.0433880	-2.8300321	4.4269019	8.8812341	-5.1594178	99.3171163	0.1293680	34.5901424	65.4343445	-1.8470825	14.7153391	-3.5426173
x7	-0.4276836	-0.4078297	-0.4025460	2.2263271	-0.5753427	0.1293680	3.0830074	15.7956453	2.5661099	0.2088372	2.9592040	0.0306690
x8	2.0584797	3.3920107	4.2083562	59.1956632	-6.3866769	34.5901424	15.7956453	720.0855371	212.7941636	-0.1252607	32.3388890	8.3821791
x9	0.4748319	-2.7854530	10.2850072	27.2410334	-1.8014211	65.4343445	2.5661099	212.7941636	784.8756534	-6.3245693	18.3962779	15.9118262
x10	0.2715656	0.3754098	-0.3127128	0.4022398	-0.1024584	-1.8470825	0.2088372	-0.1252607	-6.3245693	1.1376821	1.1452595	-0.0239257
x11	0.8262617	2.1428085	1.8419917	15.1209416	-3.0137733	14.7153391	2.9592040	32.3388890	18.3962779	1.1452595	142.2788816	16.1688301
x12	2.0976568	2.6449184	1.1598847	-1.0983981	0.4922731	-3.5426173	0.0306690	8.3821791	15.9118262	-0.0239257	16.1688301	26.0502580

Pearson Correlation Coefficients, N = 1731												
	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12
x1	1.00000	0.65402	0.05515	0.00792	-0.01825	-0.08195	-0.09735	0.03066	0.00677	0.10176	0.02769	0.16426
x2	0.65402	1.00000	0.00904	-0.00605	-0.09020	-0.08515	-0.06964	0.03790	-0.02981	0.10553	0.05386	0.15538
x3	0.05515	0.00904	1.00000	0.34131	0.17396	0.32229	-0.16634	0.11378	0.26635	-0.21271	0.11204	0.16488
x4	0.00792	-0.00605	0.34131	1.00000	-0.20032	0.08520	0.12122	0.21090	0.09296	0.03605	0.12120	-0.02057
x5	-0.01825	-0.09020	0.17396	-0.20032	1.00000	-0.38711	-0.24501	-0.17796	-0.04808	-0.07183	-0.18893	0.07212
x6	-0.08195	-0.08515	0.32229	0.08520	-0.38711	1.00000	0.00739	0.12934	0.23437	-0.17377	0.12379	-0.06965
x7	-0.09735	-0.06964	-0.16634	0.12122	-0.24501	0.00739	1.00000	0.33524	0.05217	0.11151	0.14129	0.00342
x8	0.03066	0.03790	0.11378	0.21090	-0.17796	0.12934	0.33524	1.00000	0.28305	-0.00438	0.10103	0.06120
x9	0.00677	-0.02981	0.26635	0.09296	-0.04808	0.23437	0.05217	0.28305	1.00000	-0.21165	0.05505	0.11128
x10	0.10176	0.10553	-0.21271	0.03605	-0.07183	-0.17377	0.11151	-0.00438	-0.21165	1.00000	0.09002	-0.00439
x11	0.02769	0.05386	0.11204	0.12120	-0.18893	0.12379	0.14129	0.10103	0.05505	0.09002	1.00000	0.26558
x12	0.16426	0.15538	0.16488	-0.02057	0.07212	-0.06965	0.00342	0.06120	0.11128	-0.00439	0.26558	1.00000

For Population 2

Covariance Matrix, DF = 1714												
	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12
x1	6.4006042	5.5359005	0.3696640	0.5518641	-0.0638991	-2.4206585	-0.6709136	0.1843215	-4.7629146	0.3930577	2.0195639	2.7235815
x2	5.5359005	11.0344241	0.2745043	-0.3906501	-0.2199388	-3.6155958	-0.3178403	6.8114975	-3.8556205	0.3474989	2.7213410	3.2317852
x3	0.3696640	0.2745043	2.0200633	4.5366634	0.3030599	3.7936898	-0.4259303	2.6728873	9.1815787	-0.3020790	2.7111400	1.2510814
x4	0.5518641	-0.3906501	4.5366634	107.0581947	-3.9627062	10.5262809	2.9140960	59.4458454	31.2097462	0.0462984	12.8810634	-2.6653075
x5	-0.0638991	-0.2199388	0.3030599	-3.9627062	1.9324684	-5.3768186	-0.6434049	-7.8468813	-2.2423891	-0.0332529	-1.7766228	0.3814760
x6	-2.4206585	-3.6155958	3.7936898	10.5262809	-5.3768186	99.4488354	0.7158292	44.0033829	71.8972046	-2.1140667	13.0302487	-2.9603237
x7	-0.6709136	-0.3178403	-0.4259303	2.9140960	-0.6434049	0.7158292	3.0696870	15.8911615	2.2956872	0.0994175	2.7522940	-0.0148229
x8	0.1843215	6.8114975	2.6728873	59.4458454	-7.8468813	44.0033829	15.8911615	751.8794547	176.0007426	-1.4252624	10.7988451	2.5656031
x9	-4.7629146	-3.8556205	9.1815787	31.2097462	-2.2423891	71.8972046	2.2956872	176.0007426	800.1681654	-8.8561499	15.3773124	13.3809693
x10	0.3930577	0.3474989	-0.3020790	0.0462984	-0.0332529	-2.1140667	0.0994175	-1.4252624	-8.8561499	1.1790811	0.7930623	0.2715750
x11	2.0195639	2.7213410	2.7111400	12.8810634	-1.7766228	13.0302487	2.7522940	10.7988451	15.3773124	0.7930623	137.6833602	15.4085860
x12	2.7235815	3.2317852	1.2510814	-2.6653075	0.3814760	-2.9603237	-0.0148229	2.5656031	13.3809693	0.2715750	15.4085860	24.9167263

Pearson Correlation Coefficients, N = 1715												
	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12
x1	1.00000	0.65872	0.10280	0.02108	-0.01817	-0.09595	-0.15136	0.00266	-0.06655	0.14308	0.06803	0.21567
x2	0.65872	1.00000	0.05814	-0.01137	-0.04763	-0.10915	-0.05461	0.07478	-0.04103	0.09634	0.06982	0.19490
x3	0.10280	0.05814	1.00000	0.30849	0.15339	0.26766	-0.17104	0.06858	0.22837	-0.19573	0.16257	0.17634
x4	0.02108	-0.01137	0.30849	1.00000	-0.27550	0.10202	0.16075	0.20953	0.10663	0.00412	0.10610	-0.05161
x5	-0.01817	-0.04763	0.15339	-0.27550	1.00000	-0.38785	-0.26417	-0.20586	-0.05702	-0.02203	-0.10892	0.05498
x6	-0.09595	-0.10915	0.26766	0.10202	-0.38785	1.00000	0.04097	0.16092	0.25487	-0.19523	0.11136	-0.05947
x7	-0.15136	-0.05461	-0.17104	0.16075	-0.26417	0.04097	1.00000	0.33078	0.04632	0.05226	0.13388	-0.00169
x8	0.00266	0.07478	0.06858	0.20953	-0.20586	0.16092	0.33078	1.00000	0.22691	-0.04787	0.03356	0.01874
x9	-0.06655	-0.04103	0.22837	0.10663	-0.05702	0.25487	0.04632	0.22691	1.00000	-0.28833	0.04633	0.09477
x10	0.14308	0.09634	-0.19573	0.00412	-0.02203	-0.19523	0.05226	-0.04787	-0.28833	1.00000	0.06224	0.05010
x11	0.06803	0.06982	0.16257	0.10610	-0.10892	0.11136	0.13388	0.03356	0.04633	0.06224	1.00000	0.26307
x12	0.21567	0.19490	0.17634	-0.05161	0.05498	-0.05947	-0.00169	0.01874	0.09477	0.05010	0.26307	1.00000

References:

Johnson, Richard A., and Dean W. Wichern. *Applied Multivariate Statistical Analysis*. 6th ed, Pearson Prentice Hall, 2019.

github.github.com/naity/deepufc2.Accessed 13 Dec. 2019.