**Github README: UFC Fight Analysis**

This project is a part of the AAI-500 course in the Applied Artificial Intelligence Program at the University of San Diego (USD).

Project Status: Completed

**Installation**

To use this project, first clone the repo on your device using the command below:

git init

git clone calmdownkarm/ufcdataset/versions/5

pip install -r requirements.txt

**Project Intro/Objective**

The main purpose of this project is to look at what important factors play a part in the success of a fighter. The goal of this project is to see what specific attributes of a fighter influence performance metrics. This analysis can be used for predicting fighter outcomes, used for coaching and scouting purposes. We looked at attributes pertaining to a person’s physique, their nationality, and training purposes, as well as win methods. We predict this can be used to understand possible fight outcomes and training techniques.

**Partner(s)/Contributor(s)**

·  Ashley, Andrea, Rogelio

**Methods Used**

·  Inferential Statistics

·  EDA

·  Data Visualization

·  Data Manipulation

·  Machine Learning

·  Feature Engineering

·  Correlation Analysis

**Technologies**

·  Python

**Project Description**

·   We looked at the attributes of fighters in relation to their success as fighters. We got the dataset from Kaggle. It goes from 2013 to 2017 of fighter records. We looked at height, age, weight, methods used during fighting, nationality, win type, and success of the fight.

Questions we analyzed: descriptive and inferential analysis of how various fighter features vs winning/losing outcomes - features include age, height, weight, methods used (kicks, punches, etc.). UFC fighters' nationality and training locations. Does age affect a fighter’s preferred win method (KO, SUB, or DEC)? Which factor, height or weight, has a stronger effect on a fighter’s win method (KO/TKO, Submission, or Decision)?

The hypothesis that we were exploring was to see if any of these attributes could be used to make predictions about their success. We cleaned the data, used EDA, and performed visualization. We used correlation and inferential statistics testing. We used comparisons of attributes of each fighter with their wins and losses. We used learning models to best predict fight outcomes. Some difficulties were cleaning up the data in regard to location, balancing the dataset, and making sure that we were using the best model for making each different prediction.

**License**

There are no licenses. This was created for a final project at the USD AAI-500 class, Fall 2025.

**Acknowledgments**

·  Kaggle for providing the dataset

·  Leonid Shpaner for teaching us coding techniques and analysis skills