

**04.02 정기 OMS:  
나랑 MMA 한판 뜨실 분  
619호로**

01

4월 예정 행사

## 예정 행사

01. 게임 대회 (4월 5일)

같이 게임을 하면서 친해져봐요!

02. 중간고사 시열스

다 같이 A+ 받아봅시다!!!

03. MT

우리도 MT 갑니다!!!

02

MT 컨텐츠 추천

**MT**

MT때 먹고 싶은거나 하고 싶은거 있나요?

OMS 온 사람들의 특권

03

코테 스터디 모집 중

코테 스터디

자세한 내용은 구글폼 참고!

# 04

OMS : 이준형 학우님  
나랑 MMA 한판 뜨실 분 619호로:  
데이터 분석으로 UFC 승률 예측하기

나의 은밀한 취미



## 나의 은밀한 취미



## 재밌는 OMS



MMA란?

Mixed Martial Arts

종합 격투기

복싱, 유도, 레슬링 다 됨 (근데 눈 찌르기 이런건 안됨)

**UFC? MMA?**

**UFC != MMA**

**UFC는 진짜 유명한  
MMA 단체 이름임**

**UFC? MMA?**

**나 UFC 배워! ==**

**나 프리미어 리그 배워! ==**

**나 MLB 배워!**

**UFC** 얼티밋 파이팅 챔피언십

# UFC는 언제 하나요?

UFC ON ESPN: EMMETT VS. MURPHY

UFC 314

UFC FIGHT NIGHT: HILL VS. ROUNTREE

메인 카드

예선 카드

초기 예선

4월 13일 (일) 오전 11:00  
페더급 · 메인 이벤트



알렉산더 볼카노프스키

26 - 4 - 0



디에고 로페스

26 - 6 - 0

4월 13일 (일) 오전 11:00  
라이트급 · 공동 메인



マイ클 챔들러

23 - 9 - 0



패디 펌블렛

22 - 3 - 0

4월 13일 (일) 오전 11:00  
페더급



브라이스 미첼

17 - 3 - 0



장 시우바

15 - 2 - 0

한국 표준시 기준

의견



경기 더보기

그래서 뭘 할건가요?

UFC 데이터로 승률을 예측해보자!

근데, UFC를 모르는데 승률을 예측 할 수 있을까?

## 아주 기초적인 머신러닝

생선의 무게로 이 생선이 멸치인지 참치 인지 맞춰보자!



## 아주 기초적인 머신러닝

Q. 무엇을 활용 할 수 있을까?

A. 참치와 멸치의 특성을 이용 할 수 있음

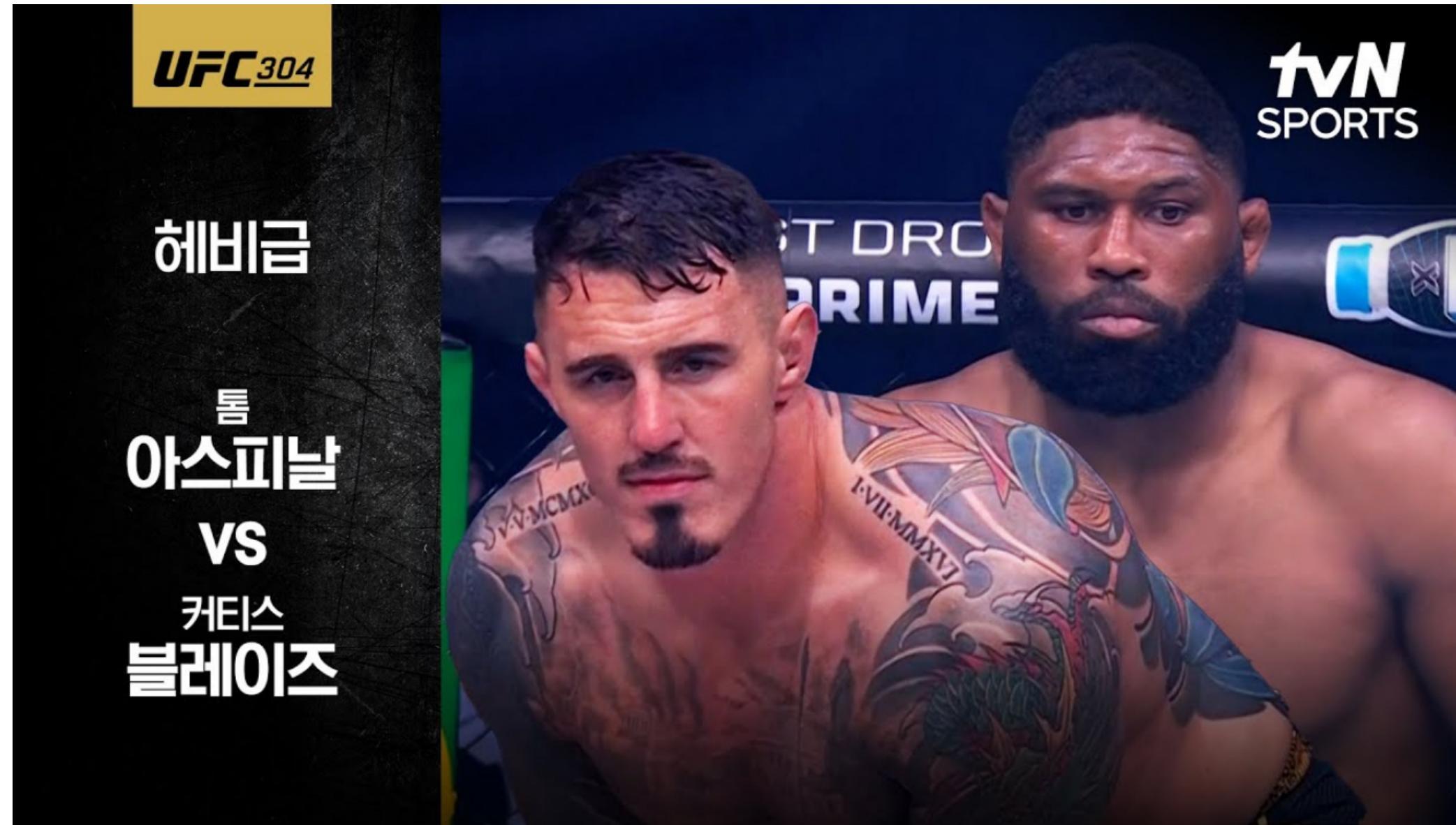
멸치는 가볍고 참치는 무거움

어떤 특성을 활용 할 수 있을까? - KO, TKO, SUB, 판정승



13분 10초

## 어떤 특성을 활용 할 수 있을까? - KO, TKO, SUB, 판정승



5분 15초

## 어떤 특성을 활용 할 수 있을까? - KO, TKO, SUB, 판정승



19분 15초

## 어떤 특성을 활용 할 수 있을까? - BJJ, 스트라이커, 레슬링

어어 형이야 ㅋ



## 어떤 특성을 활용 할 수 있을까? - 연승 기록, 연패 기록



**어떤 특성을 활용 할 수 있을까? - 신체적 특성**

**키, 몸무게, 나이, 팔 길이**

## 머신러닝이란 무엇인가요?

과거 데이터를 학습하여 패턴을 찾고,  
새로운 데이터가 주어졌을 때 결과를 예측하는 기술

## 머신러닝

생선의 무게로 이 생선이 멸치인지 참치인지 맞춰보자!

Q. 어떤 패턴이 있을까?

A. 참치는 무겁고, 멸치는 가볍다

하지만 만약에 새끼 참치라면??



## 머신러닝의 과정 - 데이터 수집

학습에 필요한 데이터를 찾는 과정

나는 kaggle을 이용하였음



## 머신러닝의 과정 - 데이터 전처리

학습에 용이하게 데이터를 바꾸는 과정

데이터가 문자열이라면?

Red : 0 / Blue : 1로 매칭

## 머신러닝의 과정 - 특성 공학

기존 변수 가공 → 새로운 변수 생성

연패 + 나이가 많음 -> 슬슬 은퇴를 생각해야



# 배당률

UFC 314: Volkanovski vs. Lopes 4월 12																	
Fighters	BetOnline	Bovada	Bookmaker	Pinnacle	Betway	MyBookie	BetUS	4casters	Bet105	SXBet	Jazz	Cloudbet	DraftKings	FanDuel	ESPN	Circa	Bet
Alex Volkanovski	-152 ▲	-140 ▲		-144 ▼	-150 —	-140 ▲	-150 ▲	-146 ▼	-144 ▼		-151 —	-142 ▼	-148 ▼	-146 ▲	-150 ▼		-150 ▼
Diego Lopes	+132 ▼	+120 ▼		+125 ▲	+120 —	+110 ▼	+126 ▼	+125 ▲	+125 ▲		+113 —	+127 ▲	+124 ▲	+114 ▼	+125 ▲		+125 ▲

## 머신러닝의 과정 - 특성 공학

타격 정확도 \* 타격 횟수

테이크 다운 횟수 \* 테이크 다운 정확도

서브 미션 횟수 \* 정확도

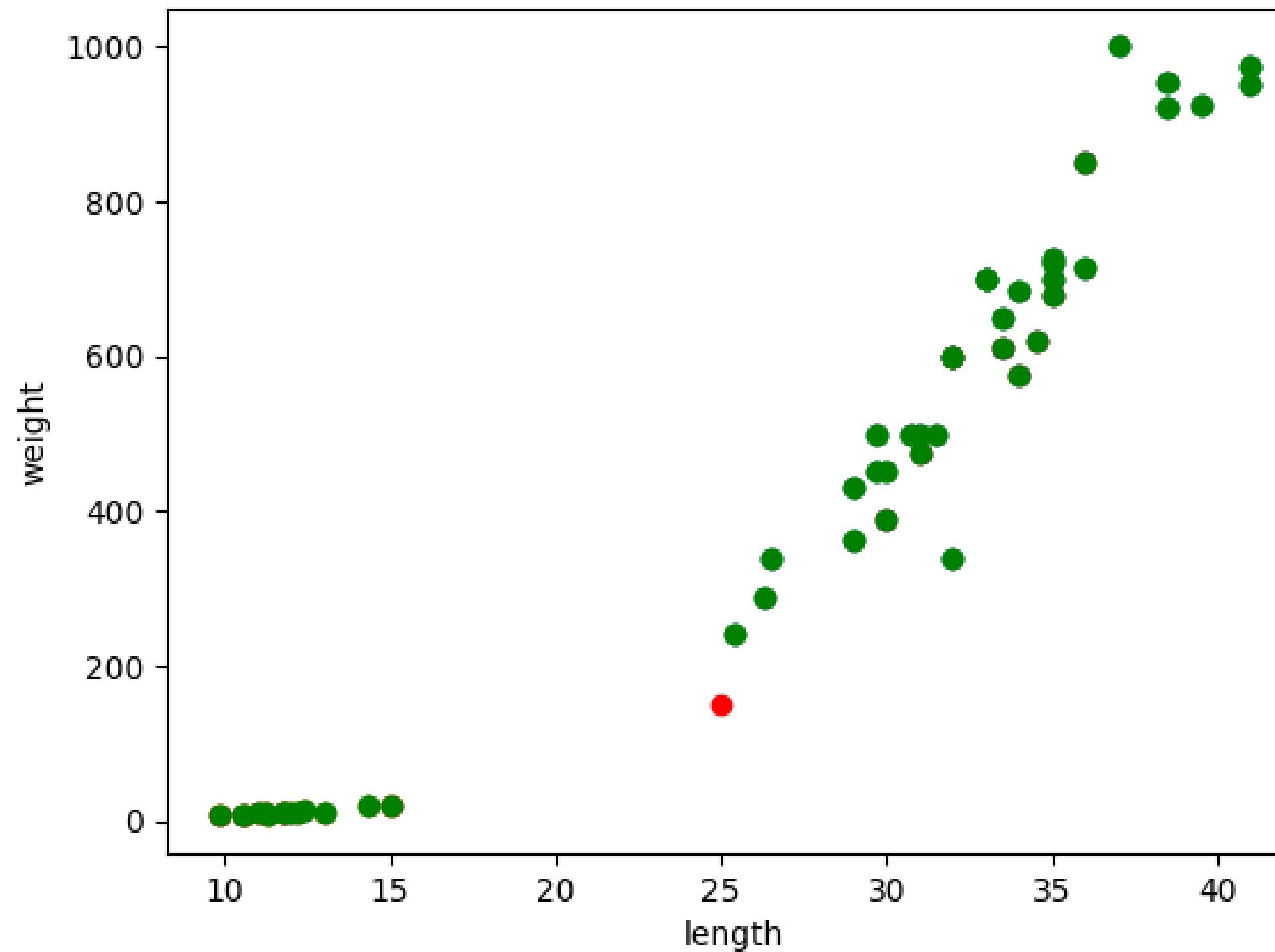
각 스타일별 상성도 파악 가능

머신러닝의 과정 - 데이터 분할

학습 데이터와 테스트 데이터 분리

모의고사 / 본 시험

## 머신러닝의 과정 - 데이터 정규화



## 머신러닝의 과정 - 모델 선택

로지스틱 회귀는 이름에 "회귀(Regression)"가 들어가지만, 실제로는 분류 (Classification) 알고리즘입니다.

회귀 모델(선형 회귀)을 기반으로 하지만, 출력을 0과 1 사이의 확률 값으로 변환하는 시그모이드 함수를 사용합니다.

## 머신러닝의 과정 - 모델 선택

### ① 선형 회귀 방정식

먼저, 선형 회귀처럼 가중치(weight)와 입력(feature)의 선형 조합을 계산합니다.

$$z = w_1x_1 + w_2x_2 + \dots + w_nx_n + b$$

여기서,

- $x_1, x_2, \dots, x_n$  : 입력 변수(features)
- $w_1, w_2, \dots, w_n$  : 가중치(weights)
- $b$  : 편향(bias)

ㅎㅇㅋㅋ



## 뭐 먹을래



## 머신러닝의 과정 - 모델 선택

### ② 시그모이드(Sigmoid) 함수 적용

로지스틱 회귀는 출력을 확률 값(0~1 사이)로 변환해야 하므로,  
시그모이드(Sigmoid) 함수를 사용합니다.

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

이 함수는  $z$  값이 어떤 숫자든 0과 1 사이의 값으로 변환합니다.

- $\sigma(z)$  값이 0.5 이상이면 클래스 1 (ex: 스팸)
- $\sigma(z)$  값이 0.5 미만이면 클래스 0 (ex: 비스팸)

결과적으로, 로지스틱 회귀의 최종 수식은:

$$P(Y = 1|X) = \frac{1}{1 + e^{-(w_1x_1 + w_2x_2 + \dots + w_nx_n + b)}}$$

즉, 주어진 입력  $X$ 에 대해 클래스 1(예: "참")일 확률을 예측하는 모델입니다.

예측을 해보자!!



# 예측을 해보자!!

```
117
118     y_proba = lr.predict_proba(X_test)[:, 1]
119     print("테스트 데이터 ROC-AUC:", roc_auc_score(y_test, y_proba))
120
121     # -----
122     # 8. 새로운 경기 데이터 예측 (예시)
123     #
124     new_fight = {
125         'RedHeightCms': 167, 'BlueHeightCms': 180,
126         'RedReachCms': 181, 'BlueReachCms': 184,
127         'RedAge': 36, 'BlueAge': 30,
128         'RedCurrentWinStreak': 0, 'BlueCurrentWinStreak': 5,
129         'RedCurrentLoseStreak': 2, 'BlueCurrentLoseStreak': 0,
130         'RedAvgSigStrLanded': 6.16, 'BlueAvgSigStrLanded': 4.20,
131         'RedAvgSigStrPct' : 0.57, 'BlueAvgSigStrPct' : 0.54,
132         'RedAvgSubAtt': 0.18, 'BlueAvgSubAtt': 2.76,
133         'RedAvgTDLanded': 1.78, 'BlueAvgTDLanded': 0.5,
134         'BetterRank': 'Red',
135         'Finish': 'Decision'
136     }
137
```

PROBLEMS    OUTPUT    DEBUG CONSOLE    TERMINAL    PORTS

```
PS C:\Users\cartc\OneDrive\문서\github\UFC> & C:/Users/cartc/AppData/Local/Microsoft/WindowsApps/python3.13.exe c:/Users/cartc/OneDrive/문서/github/UFC/ufc.py
테스트 데이터 ROC-AUC: 0.6214165487305527
C:/Users/cartc/AppData/Local/Packages/PythonSoftwareFoundation.Python.3.13_qbz5n2kfra8p0/LocalCache/local-packages/Python313/site-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LogisticRegression was fitted with feature names
  warnings.warn(
```

[새로운 경기 데이터 예측]

Red 승리 확률: 0.5846328710970039

PS C:\Users\cartc\OneDrive\문서\github\UFC>

# 배당률

UFC 314: Volkanovski vs. Lopes		4월 12																
Fighters		BetOnline	Bovada	Bookmaker	Pinnacle	Betway	MyBookie	BetUS	4casters	Bet105	SXBet	Jazz	Cloudbet	DraftKings	FanDuel	ESPN	Circa	Bet
Alex Volkanovski		-152 ▲	-140 ▲		-144 ▼	-150 —	-140 ▲	-150 ▲	-146 ▼	-144 ▼		-151 —	-142 ▼	-148 ▼	-146 ▲	-150 ▼		-15
Diego Lopes		+132 ▼	+120 ▼		+125 ▲	+120 —	+110 ▼	+126 ▼	+125 ▲	+125 ▲		+113 —	+127 ▲	+124 ▲	+114 ▼	+125 ▲		+12

## 승률로 배당률 계산하기

### 1. UFC 미국식 배당률 공식 (Moneyline Odds)

- 승률이 50% 이상일 때 (강자, Favorite)

$$\text{Moneyline} = - \left( \frac{\text{승률}}{1 - \text{승률}} \times 100 \right)$$

- 승률이 50% 이하일 때 (언더독, Underdog)

$$\text{Moneyline} = \left( \frac{1 - \text{승률}}{\text{승률}} \times 100 \right)$$

### 2. 승률 0.58 (즉, 58%)를 적용하면:

$$\text{Moneyline} = - \left( \frac{0.58}{1 - 0.58} \times 100 \right)$$

$$= -138$$



느낀점

재밌다!

## 느낀점

첫째, 특정한 한 분야에서 최고가 되는 것.

둘째, 두가지 이상의 일에서 매우 뛰어난 능력(상위 25%)를 발휘하는 것  
첫 번째 전략은 1등의 몫이다.

1등이 아닌 사람들에게는 불가능이다.

두번째 전략은 비교적 쉽다.

누구나 일정한 노력을 기울이면 상위 25퍼센트까지는 올라갈 수 있는 분야가 적어도 두 개 정도는 있다.

## 느낀점



05

뒤풀이

06

다음 연사자는요?

OMS 꽃