

A background image of a UFC fight scene, overlaid with a semi-transparent red filter. In the center, a referee is raising the arm of a victorious fighter. To the right, another fighter stands with his arms crossed. The octagon cage and various sponsor logos like 'UFCSTORE.COM' and 'P3' are visible in the background.

Команда № 4 «Python may care»

Кейс 2

Предсказание  
победителя-бойца

---

спикер:

Галахов Илья Юрьевич

команда:

Рем Алёна Владимировна

Машкина Валерия Александровна

Паисьева Ульяна Вадимовна

Антонян Анна Алексеевна

# Цель исследования

---

Создать модель, предсказывающую победителя на основе исторических статистических данных о боях лиги UFC. Модель можно будет использовать для букмекерских ставок

# ДАННЫЕ

---

## Датасет

Изначально нам было дано два датасета. Fighters и Events

## Признаки

Были отобраны следующие признаки

## Train Valid Test выборки

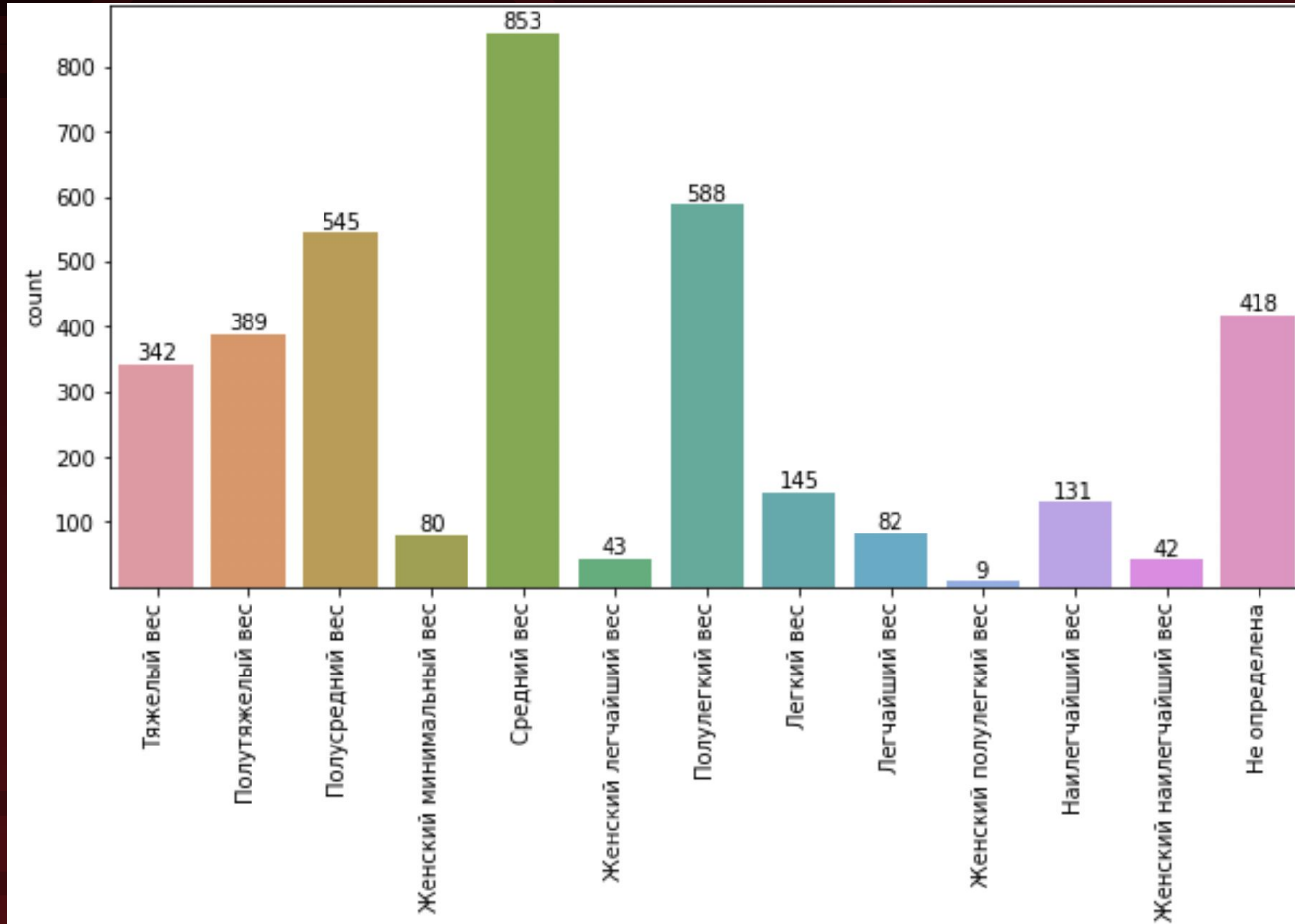
valid - 2019 год

test - 2020 год

train - все года, кроме 2019, 2020 года

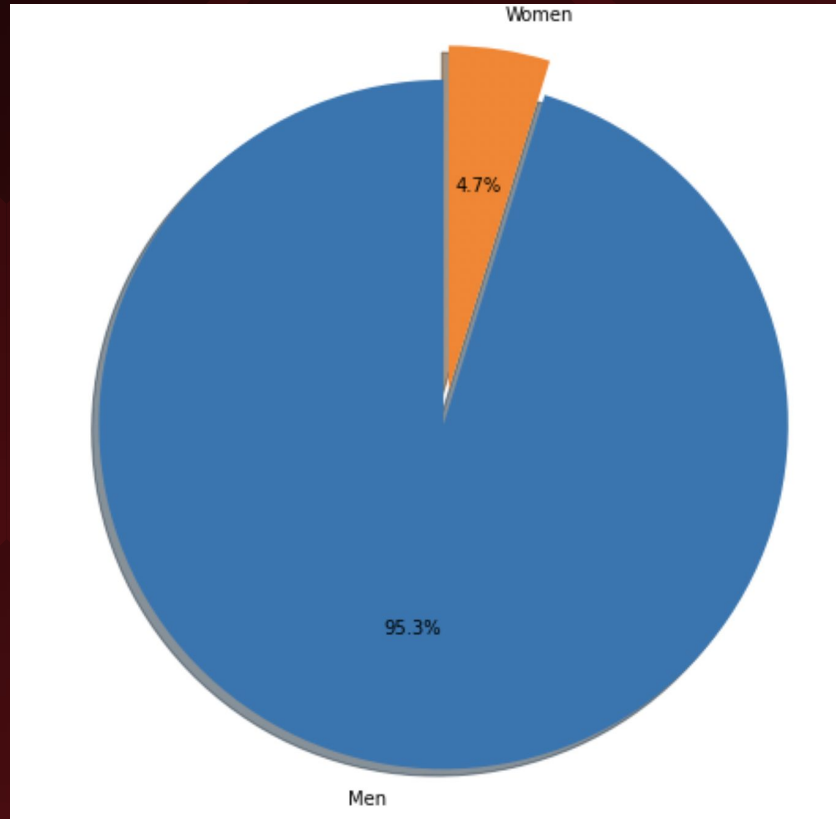
```
'duration', 'eventDate.date', 'fighterId_1', 'fighterId_2', 'id', 'rounds',  
  'weightCategory.id', "f1_hitsSuccessful", "f1_takedownTotal",  
  "f1_takedownSuccessful", "f1_submissionAttempts", "f1_takeovers",  
  "f1_accentedHitsTotal", "f2_hitsSuccessful", "f2_takedownTotal",  
  "f2_takedownSuccessful", "f2_submissionAttempts", "f2_takeovers",  
  "f2_accentedHitsTotal", "f1_weight", "f2_weight", "f1_age", "f2_age"
```

# Количество бойцов в каждой весовой категории



# Соотношение мужчин и женщин среди борцов

---



# Количество бойцов по странам

country	id
USA	1389
Brazil	304
Japan	104
Canada	86
United Kingdom	74
Russia	61
Australia	37
Mexico	28
France	27
Poland	26
Sweden	22
South Korea	20
Germany	16
Netherlands	15
China	10
Italy	9
Ireland	9
Peru	8
Iran	8
Cuba	7
New Zealand	7
Armenia	7
Argentina	7
Denmark	6
South Africa	6
Venezuela	6

Nigeria	6
Scotland	6
Spain	5
Serbia	5
Georgia	5
Jamaica	5
Philippines	5
Romania	4
Finland	4
Bosnia	4
Bulgaria	4
Democratic Republic of Congo	4
Ukraine	4
Soviet Union	4
Iraq	3
Norway	3
Lithuania	3
Colombia	3
Croatia	3
Kazakhstan	3
Czech Republic	3
Cyprus	2
Belarus	2
Belgium	2

# Посмотрим из каких стран больше всего победителей

	wins
country	
USA	17095
Brazil	4304
Japan	1824
Canada	1097
United Kingdom	991
Russia	977
Australia	410
Poland	407
France	376
Mexico	367
Netherlands	262
Sweden	245
South Korea	236
Germany	203
China	140
New Zealand	118
Argentina	109
Iran	103
Ireland	102
Italy	102
Peru	94
Soviet Union	90
Cuba	87
Armenia	84
Scotland	77

Denmark	72
Nigeria	72
Ukraine	69
South Africa	68
Jamaica	60
Serbia	58
Spain	56
Bulgaria	54
Belarus	54
Bosnia	50
Bahamas	49
Croatia	47
Cyprus	44
Iraq	42
Venezuela	39
Democratic Republic of Congo	38
Philippines	34
Finland	34
Norway	34
Romania	33
Georgia	33
Kyrgyzstan	26
Moldova	26
Floirda	25
El Salvador	25

# Гипотезы

---

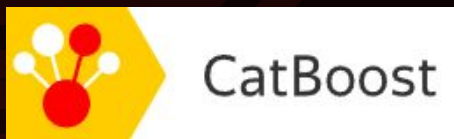
На победу может влиять:

1. Личная статистика бойца, прошлые победы и поражения
2. Параметры тела
3. Возраст
4. Национальность
5. Отличие часового пояса бойца от часового пояса проведения боя
6. Методы побед бойцов (при тренировках много времени уделяется изучению техник боя соперника, выяснение сильных и слабых сторон)
7. Кол-во определенных событий во время боя, таких как (Sub, TakeDown, КО)





# Модели



## CatBoostClassifier (CatBoost)

iterations=1000

depth=4

learning\_rate=0.01

loss\_function='Logloss'

eval\_metric='Accuracy'

F1 score = 0.72



## RandomForestClassifier (sklearn)

n\_estimators=100,

criterion='entropy',

max\_depth=10,

eval\_metric='Accuracy'

F1 score = 0.76

