



UFC FIGHT FINISH PREDICTION

CLASSIFICATION PROJECT

OUTLINE:

- Problem Statement
- Dataset
- Data Cleaning/Feature Eng
- Classification Models

Problem Statement:

- How will the fight end?

By KNOCKOUT or by DECISION or by SUBMISSION

- Goal:

To predict a fights finish



Dataset:

- 4900 Rows
- 120 Columns

	R_fighter	B_fighter	B_avg_SIG_STR_landed	B_avg_SIG_STR_pct	B_avg_SUB_ATT	B_avg_TD_landed	B_avg_TD_pct
0	Alexander Volkanovski	Brian Ortega	4.15	0.38	1.2	0.87	0.24

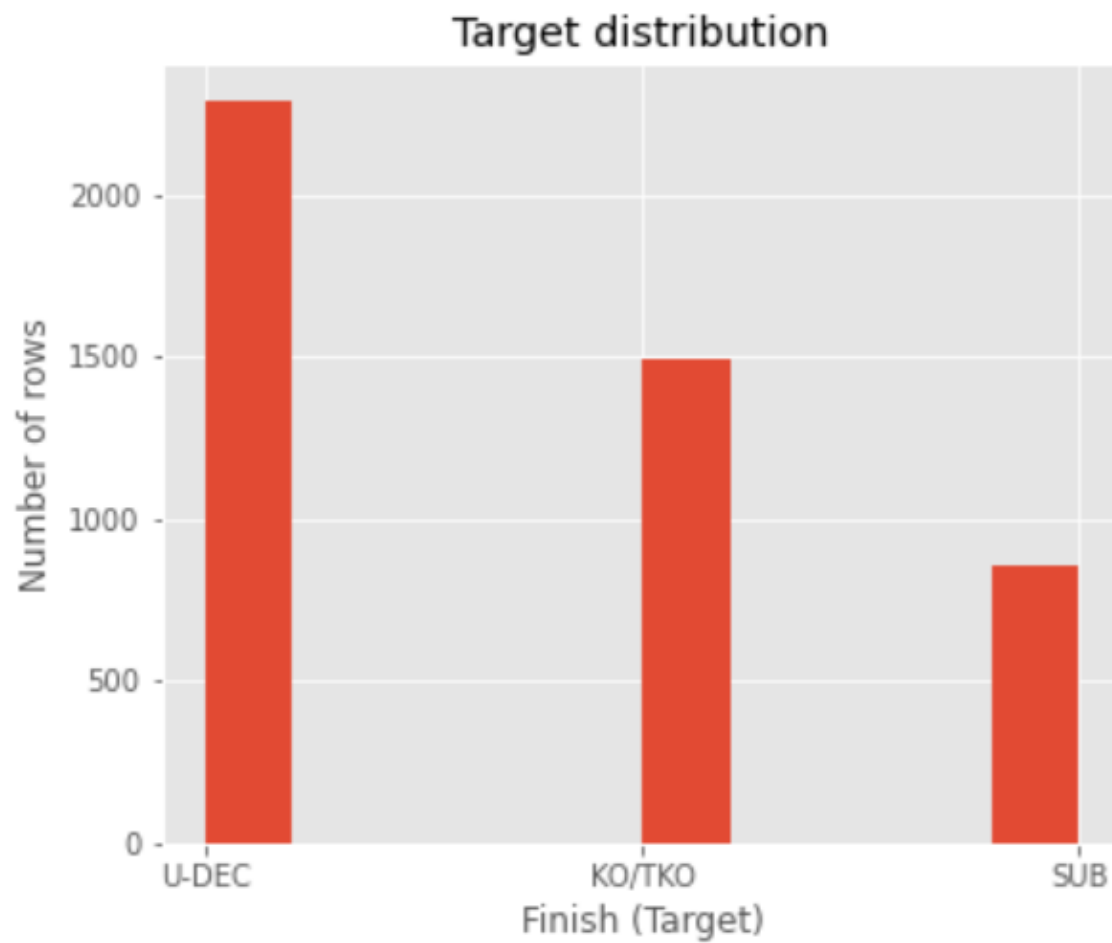
no_of_rounds	finish	finish_details	finish_round	finish_round_time	total_fight_time_secs
5	U-DEC	NaN	5.0	5:00	1500.0

Data Cleaning:

- Dropped 16 rows where target column was: **DQ, Overturned**
- Dropped 240 rows where target was **null**

Feature Engineering:

- Converting **finish round time** column to integer
- Replacing values of 0 with the column mean



Classification Models:

- Baseline Logistic Regression Model:

Training scores:

accuracy score: 0.75

f1 score: 0.63

precision score: 0.66

recall score: 0.64

Validation scores:

accuracy score: 0.78

f1 score: 0.67

precision score: 0.71

recall score: 0.67

Logistic Regression down sampling training scores:

accuracy score: 0.69

f1 score: 0.69

precision score: 0.69

recall score: 0.69

Classification Models:

- Decision Tree Model Validation Scores:

accuracy score: 0.78

f1 score: 0.71

precision score: 0.71

recall score: 0.71

- Random Forest Model Validation Scores:

accuracy score: 0.83

f1 score: 0.73

precision score: 0.77

recall score: 0.73



THANK YOU