KAZAKH-BRITISH TECHNICAL UNIVERSITY

Data mining

Final project

How reaction of crowd affect to psychology of fighter

Supervisor / Instructor: Adilet Yerkin

Students / Authors: Olzhas Koshkarbay   ID 22B030628

Nurzhan Kurmangazy   ID 22B030492

Rozimurat Mirpulatov   ID 22B031185

ALMATY KAZAHSTAN, DECEMBER 2025

Contents

1   Introduction

Combat sports are characterized not only by intense physical confrontation but also by a high level of psychological pressure. Fighters competing in the Ultimate Fighting Championship (UFC) are required to make rapid decisions, regulate emotions, and maintain concentration under extreme conditions. In addition to physical fatigue and tactical challenges, fighters are constantly exposed to external psychological stimuli, among which the reaction of the crowd plays a significant role.

Crowd reaction in UFC events can take various forms, including loud support, booing, hostile behavior, or sudden emotional silence. Such reactions may influence a fighter's psychological state by increasing stress levels, affecting confidence, provoking aggressive behavior, or, conversely, motivating higher performance. According to sports psychology theories, the presence of an audience intensifies social evaluation, which can alter emotional regulation and behavioral responses during competition.

In individual combat sports such as UFC, where a single mistake can determine the outcome of a fight, psychological stability is critical. Changes in emotional state caused by crowd reaction may lead to observable shifts in fighting style, activity level, tempo, and risk-taking behavior. Despite its potential importance, the psychological influence of crowd reaction remains difficult to measure and is rarely included in quantitative performance analysis.

This study focuses on the analysis of crowd reaction as an external psychological factor and its relationship with measurable behavioral and performance indicators in UFC fights. By applying data-driven methods to observational fight data, the research seeks to contribute to a deeper understanding of how audience behavior interacts with fighters' psychological responses inside the octagon.

2   Research Problem and Objectives

Despite the growing interest in performance analytics and sports psychology, the direct psychological impact of crowd reaction on fighters' behavior remains insufficiently explored, particularly in the context of UFC fights. Existing research mainly concentrates on physical performance metrics, home advantage, or referee bias, while the influence of audience reaction on fighters' psychological state during a bout is rarely quantified. As a result, there is a lack of empirical evidence explaining how crowd-induced psychological pressure translates into measurable behavioral and performance changes.

The objective of this study is to address this research gap by analyzing the relationship

between crowd reaction and fighters' psychological behavior and performance in UFC bouts. Specifically, the study aims to:

- examine how different types and intensities of crowd reaction affect fighters' activity level and tempo;

- analyze the influence of crowd reaction on risk-taking behavior and striking effectiveness;

- investigate whether underdog fighters are more psychologically affected by crowd reaction than favored fighters;

- evaluate the relationship between crowd reaction and fight outcomes.

To achieve these objectives, the study applies data mining techniques to a manually collected observational dataset based on the analysis of UFC fight videos and official fight statistics.

3 Data Collection

The dataset used in this project was manually collected based on real UFC fights. The main objective of the data collection process was to capture not only fight outcomes and performance statistics, but also psychological and behavioral reactions of fighters to crowd pressure.

The dataset consists of two separate tables: a fighter-level table and a fight-level table. These tables were later merged using a unique fight identifier to enable joint analysis of contextual, behavioral, and performance-related features.

3.1 Fighter-Level Data

The fighter-level table represents the core of the analysis. Each row corresponds to one fighter in one fight, meaning that every fight appears exactly twice in the dataset — once for each participant.

This table includes the following groups of variables:

- **Crowd reaction features**, such as reaction type (positive, neutral, or negative) and reaction volume measured on an ordinal scale from 0 to 3.

- **Behavioral response indicators**, including changes in activity, tempo, and risk-taking behavior following crowd reactions.

- **Performance statistics**, such as strikes landed, strikes received, significant strikes, and control time.

- **Contextual attributes**, including age, weight class, and underdog status.

- **Target variable**, indicating whether the fighter won or lost the fight.

The behavioral and crowd-related features were annotated manually based on fight observations, allowing the dataset to capture psychological dynamics that are typically absent from standard fight statistics.

3.2   Fight-Level Data

The fight-level table provides contextual information about each fight, including the event name, scheduled number of rounds, finish type, and round of finish. These variables describe the structural conditions of the fight but do not depend on individual fighter behavior.

3.3   Target Variable

The target variable of the analysis is `win`, a binary indicator equal to 1 if the fighter won the fight and 0 otherwise. The distribution of the target variable is approximately balanced, making the dataset suitable for supervised classification tasks.

4   Data Preprocessing

Prior to analysis, several preprocessing steps were applied to ensure data quality, consistency, and reproducibility.

4.1   Data Cleaning

Since the dataset was manually collected and stored in CSV format, initial cleaning was required. This included:

- Removing empty and malformed rows.

- Resolving text encoding issues during data import.

- Standardizing column names by removing extra spaces and unifying naming conventions.

Figure 3.1: Logical structure of the dataset

```
UFC Event
   |
    Fight (fight_id)
         |
         Fighter Record 1
         |       Crowd Reaction
         |       |     - reaction_type
         |       |     - reaction_volume (0-3)
         |       |
         |       Psychological Indicators
         |       |     - activity_change
         |       |     - tempo_change
         |       |     - risk
         |       |
         |       Performance Statistics
         |             - strikes_landed
         |             - strikes_received
         |             - significant_strikes
         |             - control_time_sec
         |             - win
         |
         Fighter Record 2
                 Crowd Reaction
                 Psychological Indicators
                 Performance Statistics
```
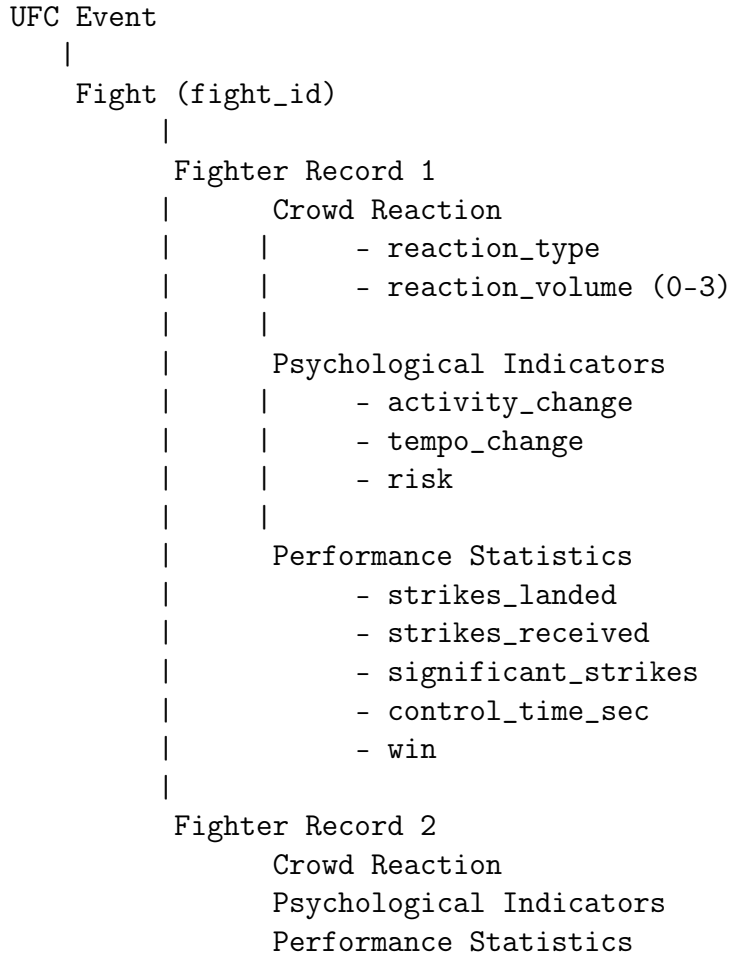
4.2 Table Merging

The fighter-level and fight-level tables were merged using the unique fight identifier. Because both tables contained a weight class attribute, suffixes were automatically added during the merge. The fighter-level weight class was retained for analysis, while the redundant fight-level column was removed.

4.3 Data Validation

Several validation checks were performed after preprocessing:

- Each fight appears exactly twice in the merged dataset.

- No missing values were found in key analytical variables.

- The target variable is balanced across classes.

These steps ensured that the dataset was structurally correct and suitable for both supervised and unsupervised learning methods.

5   Methodology

The analysis combines exploratory data analysis, supervised learning, and unsupervised learning techniques to investigate the psychological impact of crowd reaction on fighters.

5.1   Exploratory Data Analysis

Exploratory analysis was conducted to examine relationships between crowd reaction, behavioral changes, and fight outcomes. Visualizations and aggregated statistics were used to identify patterns such as:

- The association between crowd reaction intensity and behavioral instability.

- Differences in behavioral responses across age groups and weight classes.

- Variation in win probability under different crowd reaction conditions.

5.2   Supervised Learning

To quantify the effect of behavioral and performance features on fight outcomes, supervised classification models were applied. Logistic Regression was selected as the primary model due to its interpretability, while Random Forest was used as a robustness check.

Model performance was evaluated using accuracy, precision, recall, F1-score, and confusion matrices.

5.3   Unsupervised Learning

Unsupervised learning was applied using K-Means clustering to identify hidden behavioral patterns. Clustering was performed without using the target variable to avoid data leakage. The resulting clusters were interpreted as distinct behavioral response profiles under crowd pressure.

6 Results

The results demonstrate that crowd reaction influences fight outcomes primarily through behavioral changes rather than direct effects.

6.1 Supervised Learning Results

Logistic Regression achieved an accuracy of approximately 77%, with balanced precision and recall. Performance-related variables such as strikes received and strikes landed were the strongest predictors of fight outcome. However, behavioral variables related to crowd reaction retained explanatory power even after controlling for performance.

Random Forest produced similar performance, confirming the robustness of the findings.

Table 6.1: Feature Importance for Logistic Regression and Random Forest

Logistic Regression		Random Forest	
Feature	**Coefficient**	**Feature**	**Importance**
strikes_landed	0.817628	age	0.220134
significant_strikes	0.529912	strikes_received	0.209734
activity_change	0.264666	strikes_landed	0.145548
reaction_volume	0.130545	significant_strikes	0.136166
risk	0.115648	is_underdog	0.074083
control_time_sec	0.077540	control_time_sec	0.064820
is_underdog	-0.115275	reaction_volume	0.057558
tempo_change	-0.252949	risk	0.036059
age	-0.300283	activity_change	0.033830
strikes_received	-1.573357	tempo_change	0.022068

6.2 Clustering Results

Clustering analysis identified six distinct behavioral profiles. Clusters characterized by adaptive behavior combined with technical control exhibited the highest win rates. In contrast, passive fighters and those showing unstable tempo or excessive reaction to crowd pressure demonstrated lower success rates.
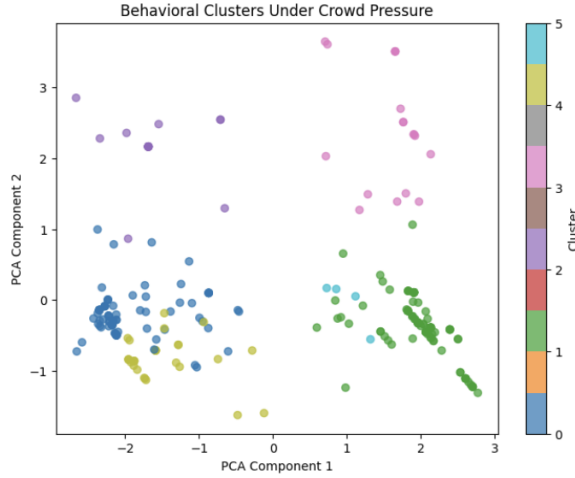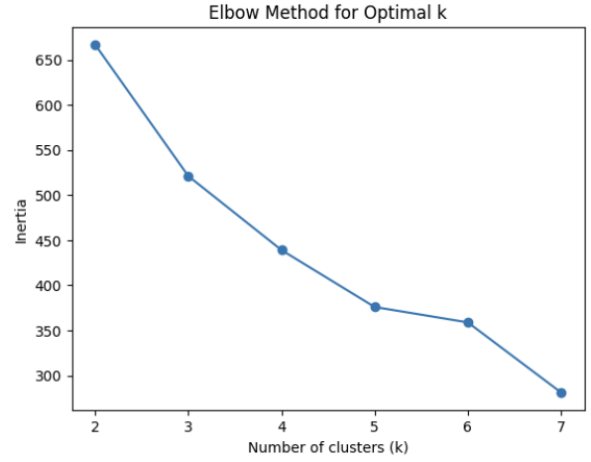
Figure 6.1: Clustering results, $k = 6$

Figure 6.2: Cluster profiles, $k = 6$

7   Visualization

Visualizations played a key role in interpreting the results. Line plots and bar charts were used to illustrate how behavioral variables vary across age groups, weight classes, and crowd reaction types. A two-dimensional PCA projection was applied to visualize the clustering structure and highlight distinct behavioral archetypes.

8   Conclusion

This study shows that crowd reaction affects fight outcomes indirectly by influencing fighter behavior. Behavioral responses such as tempo disruption and excessive risk-taking are associated with lower probabilities of winning, while adaptive behavior combined with technical control leads to higher success rates.

The findings highlight the importance of psychological resilience in high-pressure competitive environments and demonstrate the value of combining supervised and unsupervised learning techniques in sports analytics.

9   Related Work

The influence of audience presence on athletic performance has been widely discussed in sports psychology literature. Numerous studies have examined the phenomenon of home advantage, suggesting that crowd support can positively affect athletes' motivation and confidence while also influencing referee decisions. Research has shown that crowd

noise and audience size may alter emotional states and decision-making processes during competition.

In individual sports, psychological pressure from spectators has been linked to increased stress levels and changes in performance consistency. Some studies indicate that social evaluation caused by audience presence can lead to performance anxiety, particularly in high-stakes situations. In combat sports, where emotional regulation and concentration are essential, such pressure may have a stronger and more immediate effect.

However, existing research on combat sports primarily focuses on physical conditioning, technical efficiency, and injury prevention. Quantitative studies that directly analyze the psychological impact of crowd reaction on fighters' in-fight behavior are limited. In the context of UFC, most available analyses rely on official statistics without incorporating external psychological factors such as audience behavior.

This study differs from previous work by explicitly incorporating crowd reaction as a measurable variable and linking it to behavioral and performance indicators. By using a manually constructed dataset based on direct observation of UFC fights, the research provides a novel data-driven perspective on the psychological dynamics of professional combat sports.

References

[1] ESPN. UFC News, Results, and Schedule.

[2] Sherdog. Mixed Martial Arts Fights, Events, and Fighter Database.

[3] UFC Stats. Official Fighter Statistics and Records.