UVIC
UNIVERSITAT DE VIC
UNIVERSITAT CENTRAL
DE CATALUNYA

University of Vic

Impact of Fan Engagement on Football Game Outcomes: A Data-Driven Analysis

Eliott Coutaz, Christian Saidi, Clement-Louis Joyat
University of Vic, Spain

INTRODUCTION

Football's Global Fascination

Football, a sport revered globally, captivates millions with its blend of skill, strategy, and endurance. Yet, the role of fans, especially their social media engagement, in influencing match outcomes remains underexplored. This study delves into how fans' online interactions impact football games, focusing on the 2022 World Cup and Morocco's team.

Impact of Fan Engagement

Central to our research is quantifying fans' online engagement effect on team performances. By analyzing Twitter interactions, we assess if fans' emotions correlate with football match outcomes. Our primary case study is the high-stakes environment of the 2022 World Cup, showcasing a new perspective on sports, social media, and collective emotions.

METHODS

- **Data Collection** - Utilized Twitter API and sports data API for tweets and game results, focusing on the 2022 and 2018 World Cup matches.
- **Data Processing** - Employed TextBlob for sentiment analysis and Pandas for data structuring, ensuring comprehensive data handling.

Sentiment Analysis:

Tweets related to the 2022 World Cup were classified into positive, negative, or neutral categories using TextBlob. We mapped these sentiment distributions against match outcomes to identify any correlations, providing an innovative lens to understand fan influence on games.

Predictive Modeling:

We constructed a machine learning model to investigate the potential impact of fan sentiment on game results. This model integrated sentiment analysis data with game statistics, aiming to predict match outcomes based on RandomeForest Classification method.

Comparative Analysis:

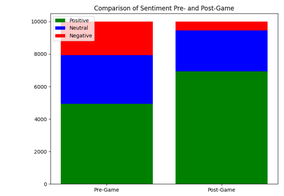
Our approach involved a detailed comparison of fan sentiments before and after matches. By examining these shifts, we gained insights into the relationship between team performance and the emotional responses of fans, highlighting the dynamic nature of this interaction.

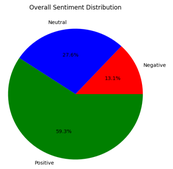
Model Evaluation:

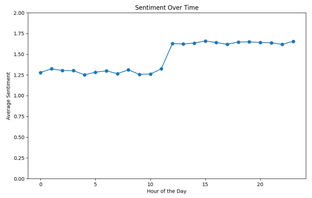
To ensure the reliability of our findings, the predictive model underwent rigorous evaluation. We used standard classification metrics and Receiver Operating Characteristic (ROC) curves to assess the accuracy and predictive strength of the model, validating its effectiveness in linking fan sentiment with football game outcomes.

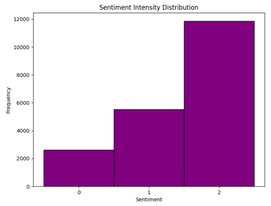
RESULTS

SENTIMENT ANALYSIS OF THE 2022 MOROCCO VS SPAIN

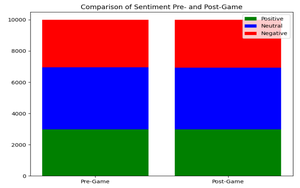


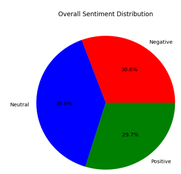


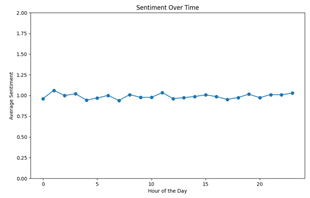


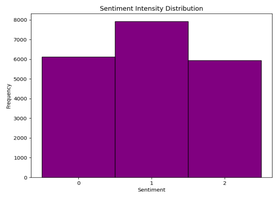


SENTIMENT ANALYSIS OF THE 2018 MOROCCO VS SPAIN





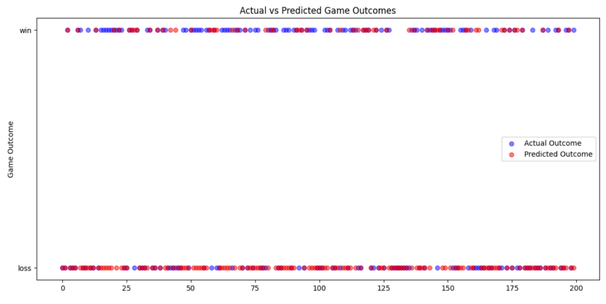


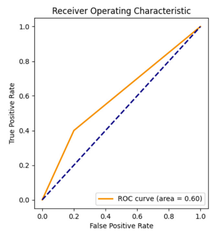


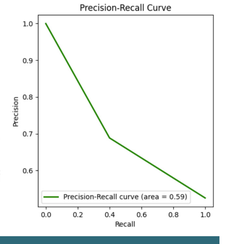
PREDICTIVE MODELING FOR CORRELATING FAN SENTIMENT AND GAME OUTCOMES

Objective: To predict football match outcomes based on Twitter fan sentiment, we developed a machine learning model using the RandomForest classifier. The model was trained on a dataset comprising sentiment texts and corresponding game outcomes, utilizing TF-IDF for feature extraction and a RandomForest classifier for prediction.

Class	Precision	Recall	F1-Score	Support
Loss	0.55	0.80	0.65	95
Win	0.69	0.40	0.51	105
Accuracy			0.59	200
Macro Avg	0.62	0.60	0.58	200
Weighted Avg	0.62	0.59	0.57	200







Prediction: Each point on the graph represents a game, with the horizontal axis numbering the games and the vertical axis denoting the outcome (win or loss). The proximity of the predicted outcomes (red) to the actual outcomes (blue) illustrates the model's accuracy on a game-by-game basis.
Evaluation: The model's performance was further visualized using ROC and Precision-Recall curves, with the ROC curve demonstrating a 60% chance of correctly distinguishing between classes (AUC = 0.60) as seen on 10, and the PR curve indicating moderate performance (area = 0.59).

REFERENCES

1. Coutaz, E., Final Project Data Science. GitHub repository for this paper. Available at: https://github.com/GnomeGuerrier/Final-project_Data_Science.git.
2. Singh, T., Twitter Sentiment Extraction - Analysis, EDA and Model, Kaggle (2020). Available at: https://www.kaggle.com/code/tanulsingh077/twittersentiment-extaction-analysis-eda-and-model.
3. Entity Sentiment Analysis, Kaggle Dataset. Available at: https://www.kaggle.com/datasets/jp797498e/twitter-entity-sentiment-analysis.
4. Biocomputing-Teaching, Data Science with Python Course Materials, GitHub repository. Available at: https://github.com/Biocomputing-Teaching/DataScience-with-Python.v
5. Pal, S., Awesome Twitter Data. GitHub repository. Available at: https://github.com/shaypal5/awesome-twitter-data.

ACKNOWLEDGMENTS

We thank Professor Jordi Villà Freixa for guiding our research. His insightful questions and expertise greatly shaped our analysis and understanding, enriching our work and tool utilization.

Carrer de la Sagrada Família, 7, 08500 Vic, Barcelona
www.uvic.cat
uvic@gmail.com

Morocco vs Spain 2022 - Image of the moroccavn team celebrating after the end of the game.

DISCUSSION

Exploring Digital Influence in Sports

This study highlights the potential influence of fan sentiment on social media over sports outcomes, exemplified by the correlation between positive Twitter engagement and Morocco's World Cup performance. While insightful, our findings are limited by the data scope and sentiment analysis biases. These aspects emphasize the need for careful interpretation and suggest areas for methodological improvements in future research.

Future research should broaden the scope to various sports and social media platforms, integrating cultural factors into sentiment analysis. This could unveil more complex relationships between digital fan interaction and team performance, offering valuable insights for sports analytics and team strategies in the digital era. The intersection of fan engagement, sports success, and data science opens new avenues for exploration in understanding the digital influence on sports.

CONCLUSION

This study uncovers the nuanced relationship between fan sentiment on social media and football match outcomes, with a focus on the Moroccan team's performances in the 2018 and 2022 World Cups.

Our analysis revealed a tangible correlation between the positive fan engagement, particularly in 2022, and the team's unexpected success, challenging the usual paradigms of sports performance analysis. While our predictive model showed moderate accuracy, it highlights the potential influence of emotional support on sports teams.

These findings invite further exploration into the role of social media sentiment in the dynamics of competitive sports, suggesting that fan engagement might be an underrecognized factor contributing to a team's success.