Predictive Analysis of Player Performance and Game Outcomes in CS:GO Matches

Kiran Kumar Nimmakayala Sam Miller Aadarsh Eetikala

1. Summary

In the world of e-sports, Counter-Strike: Global Offensive (CS:GO) is a forerunner of the first-person shooter genre. Beyond being just a game, CS:GO incorporates a complex mix of strategy, mechanical ability, and mental agility. Our project dives into this complex world with the goal of unraveling insights hidden within data extracted from over 1400 competitive matches. Through careful analysis and the application of advanced machine learning techniques, we seek to uncover recurring patterns that can be used to predict the outcomes of a match and an individual player's performance. This project is about elevating the competitive spirit of CS:GO by offering new, data-driven insights that can improve training, decision-making, and gameplay strategies.

2. Problem Overview

Our primary objective is to extract meaningful insights from an extensive dataset of CS: GO matches, with a specific focus on predicting player performance and the outcomes of matches. In a game where quick decisions and strategies play a crucial role, understanding patterns in player behavior and game dynamics is critical. We are using machine learning models to analyze data from over 1400 matches to make these predictions. The goal is to provide players and coaches with useful insights that can enhance strategy, improve skills, and increase the likelihood of winning.

A secondary goal of our project is to ascertain the best machine-learning technique for predicting player performance and match outcomes. This way, we can identify the methods that produce the most accurate and reliable predictions.

3. Data Description

We have sourced a comprehensive dataset from Kaggle that provides a granular view of over 1400 competitive CS:GO matches and captures a wide array of in-game events and dynamics. Features of this dataset include data detailing player interactions, such as throwing grenades, dealing damage, and kills/deaths. In the dataset, this information is individually tabulated, with one row corresponding to one action. Each row also contains statistics about the interaction, such as server tick ID, seconds into the round, weapon type, and the identities of the attacker and victim. The dataset also contains higher-level information about each round, including the victors/losers, the time elapsed, the map, and the round identifier.

The point of our analysis revolves around two primary target variables: individual player performance, measured through metrics like kills and damage, and the overall game outcome, identifying the winning side. This dataset forms the bedrock of our project, setting the stage for our predictive analysis endeavors.

4. Analysis Plan

4.1 Data Preprocessing and Feature Selection

Our first step involves cleaning and preprocessing the dataset to ensure accuracy and reliability in the subsequent analysis. We'll handle missing values, outliers, and potential inconsistencies to create a clean and robust dataset. Feature selection will be pivotal; we'll employ statistical methods and domain knowledge to identify, retain, and transform features that are significant in influencing game outcomes and player performance. Techniques like correlation analysis and feature importance ranking will assist in refining the dataset, ensuring it's optimized for training effective machine learning models.

4.2 Model Evaluation

With a refined dataset, we will train various machine-learning models, including logistic regression, decision trees, random forests, and gradient-boosting machines, among others. Each model will be tailored and optimized to capture the intricate patterns within the CS:GO data effectively. Hyperparameter tuning will be executed to enhance model performance, ensuring precision in predictions.

Evaluation metrics are instrumental in assessing model effectiveness. We'll employ a combination of accuracy, precision, recall, F1-score, and AUC-ROC to gauge the models' performance in predicting player performance and game outcomes. Cross-validation strategies will be incorporated to validate the models' effectiveness and reliability on unseen data.

5. Goal and Timeline

Our primary goal is to develop a predictive model that accurately forecasts player performance and game outcomes in CS:GO, transforming raw data into actionable insights that can refine training and in-game strategies. The timeline for achieving this is outlined below:

Timeframe	Tasks
Week 1-2	▪ Data pre-processing to clean and optimize the dataset. ▪ Conduct an exploratory analysis to uncover initial insights and identify key features.
Week 3-5	▪ Build and train various machine learning models. ▪ Experiment with different algorithms, tuning parameters, and features to optimize predictions. ▪ Evaluate and compare model performance. ▪ Fine-tune and ensemble top-performing models to enhance predictive accuracy.
Week 6	▪ Compile insights and findings into a comprehensive final report, highlighting methodologies, results, and practical applications of the insights.

References:

1. CS:GO Competitive Matchmaking Data. Retrieved from Kaggle Dataset
2. Bahrololloomi, F., Klonowski, F., Sauer, S. *et al.* E-Sports Player Performance Metrics for Predicting the Outcome of League of Legends Matches Considering Player Roles. *SN COMPUT. SCI.* **4**, 238 (2023). https://doi.org/10.1007/s42979-022-01660-6
3. W. Xu Huang, J. Wang and Y. Xu, "Predicting Round Result in Counter-Strike: Global Offensive Using Machine Learning," *2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP)*, Xi'an, China, 2022, pp. 1685-1691, doi: 10.1109/ICSP54964.2022.9778597.