Valorant S-Tier Team Winner Prediction using Machine Learning

by Mayank Bharti

1 Introduction

Esports analytics has emerged as a powerful application of data science, enabling teams, analysts, and organizers to gain insights into competitive performance. Valorant is a popular tactical shooter esports game that hosts tournaments of varying competitive levels. Among these, S-Tier tournaments represent the highest level of competition.

This project focuses on building a machine learning classification model to predict whether a Valorant esports team qualifies as an S-Tier winner based on historical performance metrics.

2 Problem Statement

The dataset used in this project does not explicitly label teams as winners or losers. Therefore, the objective of this project is:

> To predict whether a Valorant team is an S-Tier winner using historical performance data.

This is formulated as a binary classification problem:

- 1 – S-Tier Winner
- 0 – Non S-Tier Team

3 Dataset Description

The dataset was obtained from Kaggle and contains aggregated performance statistics of professional Valorant teams.

3.1 Features

- Gold: Number of gold medal finishes
- Silver: Number of silver medal finishes

- Bronze: Number of bronze medal finishes
- Earnings: Total prize money earned
- S Tier: Number of S-Tier tournament participations

4 Data Preprocessing

4.1 Data Cleaning

The Earnings column originally contained currency symbols and commas. These were removed and the column was converted into numeric format. Missing values were handled appropriately.

4.2 Target Variable Creation

A domain-based rule was applied:

- Teams with at least one S-Tier participation were labeled as Winners
- Teams with zero S-Tier participation were labeled as Non-Winners

5 Feature Selection

To avoid data leakage, the S Tier column was excluded from the feature set as it was used to define the target variable.

Two feature configurations were tested:

- Initial Model: Gold, Silver, Bronze, Earnings
- Final Model: Gold, Silver, Bronze

6 Machine Learning Methodology

6.1 Model Selection

Logistic Regression was chosen due to its simplicity, interpretability, and suitability for binary classification.

6.2 Feature Scaling

All numerical features were standardized using StandardScaler to ensure uniform contribution.

6.3 Handling Class Imbalance

The dataset contained significantly fewer S-Tier teams. Class-weighted logistic regression was used to address this issue.

7 Model Training

The dataset was split into 75% training data and 25% testing data using stratified sampling. The model was trained on standardized training data.

8 Model Evaluation

8.1 Evaluation Metrics

- Accuracy
- Precision
- Recall
- F1-score

8.2 Results

The final model achieved approximately 80.5% accuracy with improved recall for S-Tier teams.

9 Impact of Earnings Feature

Including Earnings resulted in nearly 99% accuracy and precision due to its strong correlation with S-Tier participation. To ensure meaningful learning, Earnings was removed from the final model.

10 Feature Importance Analysis

Feature importance was derived from logistic regression coefficients and visualized using Matplotlib.

Feature	Importance
Gold	Positive (Strongest)
Bronze	Positive (Moderate)
Silver	Negative

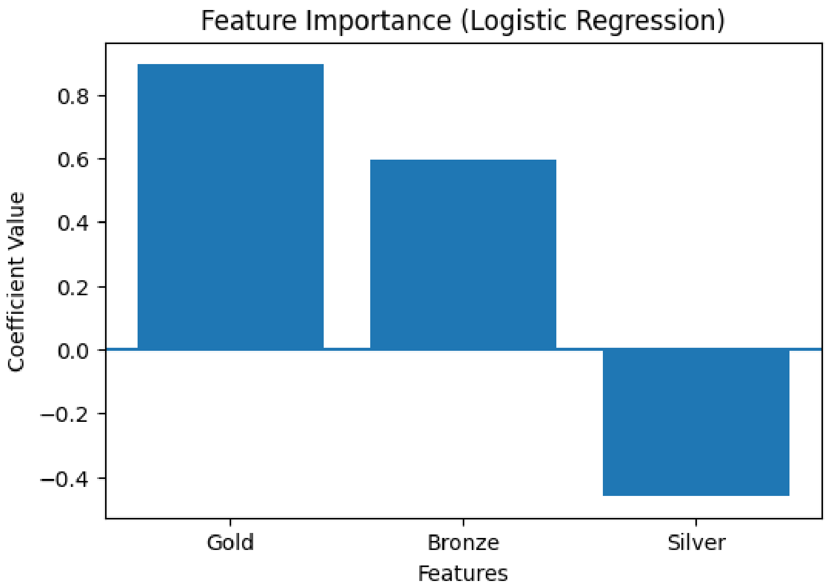

Figure 1: Features Importance

11 Prediction on New Data

The trained model was tested on unseen team statistics and produced both class predictions and probability-based confidence scores.

12 Conclusion

This project demonstrates the effective application of machine learning techniques to esports analytics by combining domain knowledge, proper preprocessing, and balanced evaluation strategies.