Project Proposal
Zacharias Johansson **|** Albatool Sassi **|** Afia Silva
21036299 **|** 23021878 **|** 23028048

Application problem

Question

"Can machine learning predict the outcome of an esports match using historical in-game statistics"

Esports has recently grown into a billion-pound industry. Accurately predicting the outcomes of matches can benefit esports teams, analysts and betting markets. Our model will be focused on Multiplayer Online Battle Arena (MOBA) games and First-Person Shooter (FPS) games. The complexity of these genres mean that there are many factors that influence matches and there can be a lot of variance in results.

Context and Background

Traditional sports prediction has been researched extensively, however esports is understudied in ML. Our model will look to extend predictive models to MOBA (League of Legends, Dota 2) and FPS (CS:GO) games by analysing match data.

Differences to regular sports:
- Player stats can change throughout the game
- Character selection and team composition can dictate the game
- Factors like map selection and team synergy can affect outcomes
- Team performance is far more important than individual performance.
- Updates to the game can completely change strategies, often referred to as the "meta".

The Dataset

Our model will be trained using publicly available datasets from Kaggle, we can also gather data from "OpenDotaAPI" and "Riot Game API" if need be.

Features

- Player stats (Kill death ratio, gold earned p/m, experience earned p/m)
- Team performance (Objective control, team gold/experience)
- Game factors (Match duration, game version, team composition)

We will pre process the data by removing outliers, normalization and encoding data into categories.

Methodology

Our approach consists of four steps: data preprocessing, model selection, training and evaluation.

We will train and compare different ML models:

- Logistic Regression
- Tree-based models
- Support Vector Machines (SVM)
- k-Nearest Neighbors (k-NN)

We will compare these models against each other and determine which is the most reliable.

Evaluation

Model performance will be evaluated using:

- Accuracy (Percentage of correct match predictions
- Precision (Measures what number of prediction wins were correct)
- Recall (Measures how well the model captures actual wins)
- F1-score (mean of precision and recall)
- Receiver Operating Characteristic (ROC) - Area Under Curve (AUC) (Measures how well the model distinguishes between wins and loses.

Different models will be evaluated and compared to each other using these metrics. By applying ML to esports predictions, this model can help teams, analysts and fans understand what drives success in competitive gaming.