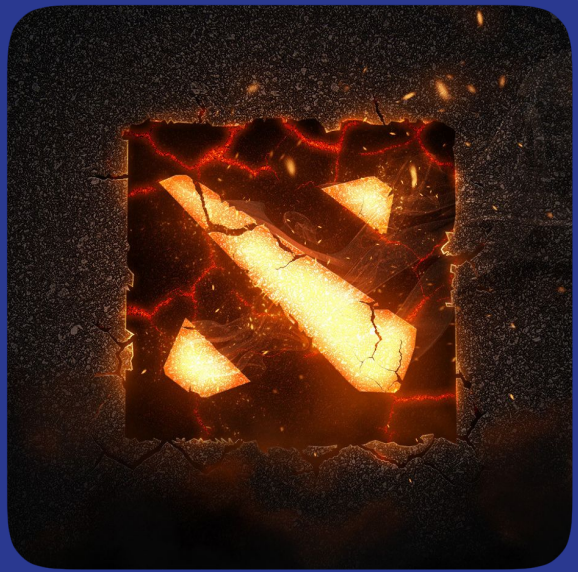
# Dissecting DOTA

Statistical analysis of Dota 2



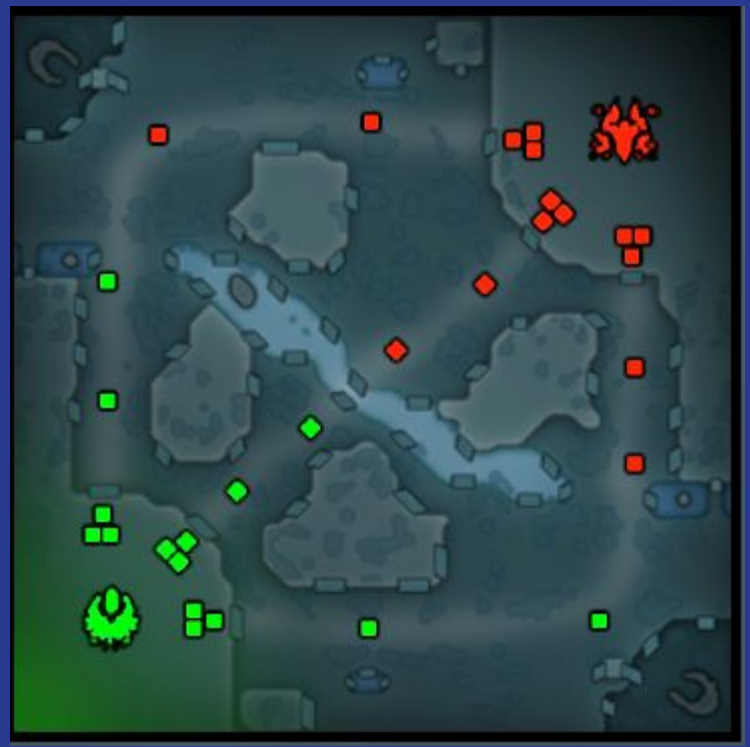
- Sathkrith (1st year MSCS)
- Madhukara S Holla (1st year MSCS)

# Background

- Game of 10 players, 5 in each team
- Each player controls a unique hero
- Objective is to destroy the enemy's Ancient (core structure)

- Different roles in the game
  - Carry
  - Support
  - Mid / off-laner
- Resources and buildings
  - Gold and Experience
  - Towers and Barracks

- Team strategies and dynamics
  - Resource acquisition
  - Item acquisition
  - Phases of the game



# Problem Description

Does a specific distribution of in-game resources among various team roles lead to a higher likelihood of winning?

What constitutes this optimal distribution strategy?

# Summary of the Data

Dataset:

- Player roles
- Heroes picked
- Player rank - skill level

- Resource metrics with timestamps
  - Gold earned
  - Experience gained
  - Player net-worth
  - Timestamp
  - Objectives secured

- Match duration
- Match outcome - win / loss

Assumptions:

- Independence: Each match is an independent event
- Homoscedastic: Variance of resource distribution is consistent across matches
- Linearity: Relationship between resources and match outcome is linear
- No perfect multicollinearity: No two variables predict each other perfectly (some correlation exists: gold vs exp)
- No extreme outliers: All players in the game are similarly skilled

Sampling restrictions:

- Limit the number of matches having the same players
- Matches should be sampled from the
  - Same skill bracket
  - Same patch of the game

# Statistical Analysis

- Proposed methodology: Multiple Logistic Regression
- Individual models according to game phase:
  - 10 minute - Early game
  - 30 minute - Mid game
  - End of the game

- Response variable: Match outcome

- Predictors:
  - Resources for each role
  - Player roles
  - Overall team resources
  - Interaction terms - Player x Resource

- Additional variables to consider:
  - Objectives captured
  - Game duration
  - Enemy team's resources

Difficulties:

- Interaction and correlations
  - Resources are not isolated
  - Multicollinearity due to player interaction and synergies

- Random effects
  - Different heroes and playstyle
  - Different game strategies

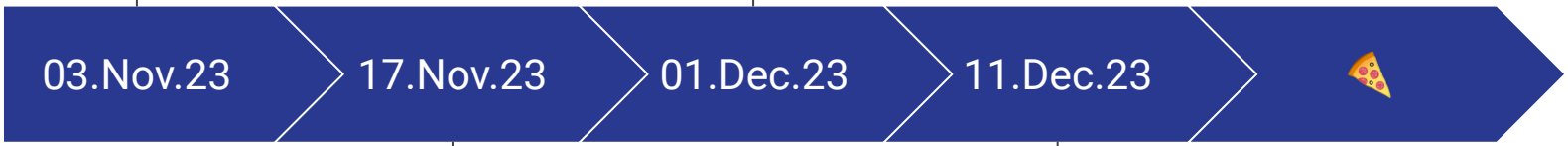
- Correlation vs Causation
  - Resource allocation has an association with the outcome, but it doesn't mean a specific allocation increases chances of winning

Finish EDA and Confirm model assumptions

Formulate Hypothesis

Finalize methodology

Finish Hypothesis Testing and Validation




Finish statistical analysis
Explore results

Report Submission

# Questions

# Thank You