
DOTA2
SQUARE 2011 SQUAR3X.DEVIANTART.COM

Dataset and Context

- 50.000 games of Dota2
- 18 different .csv files

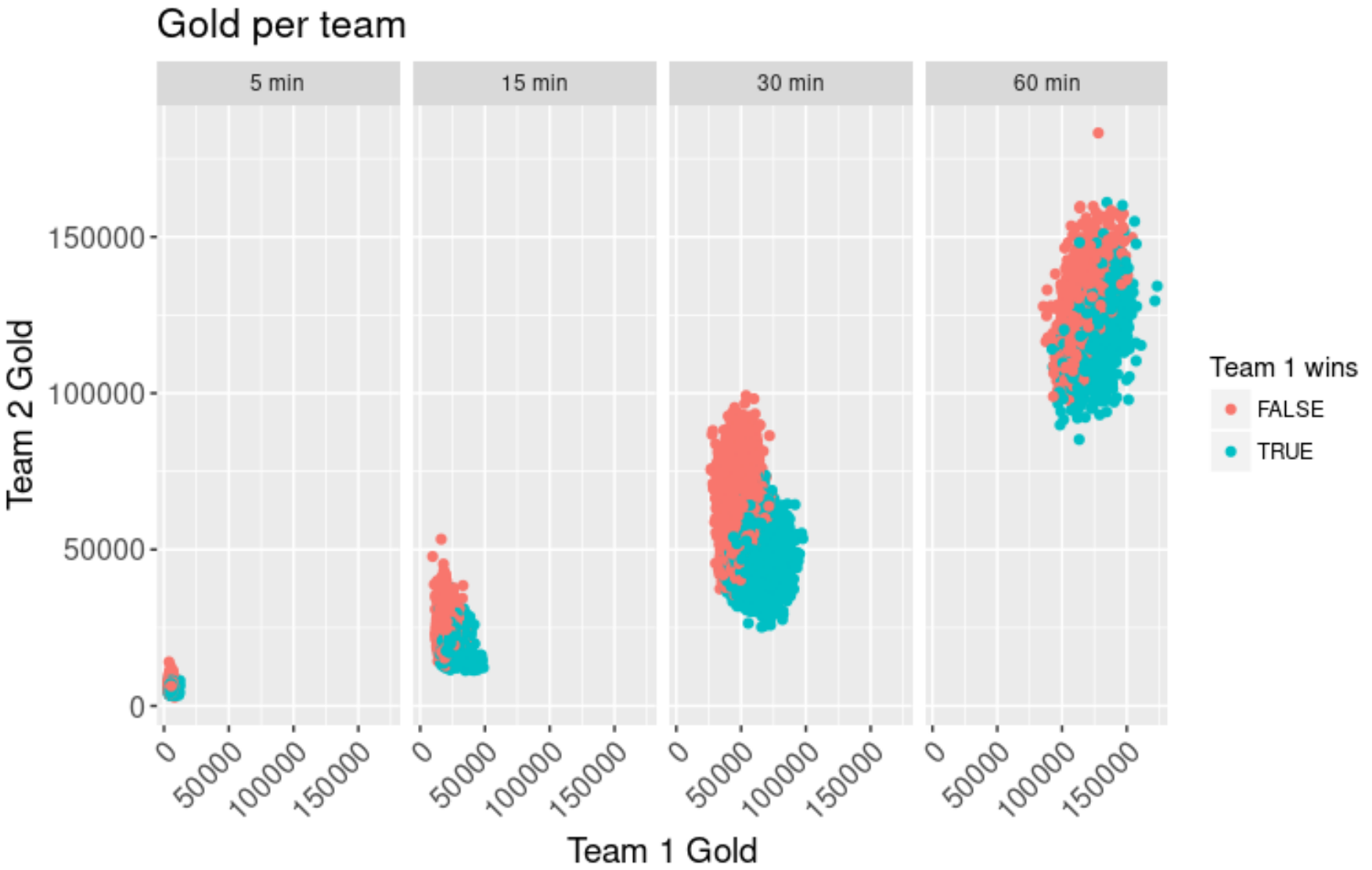
- Two types of data:
  - Match summaries
  - In-game data

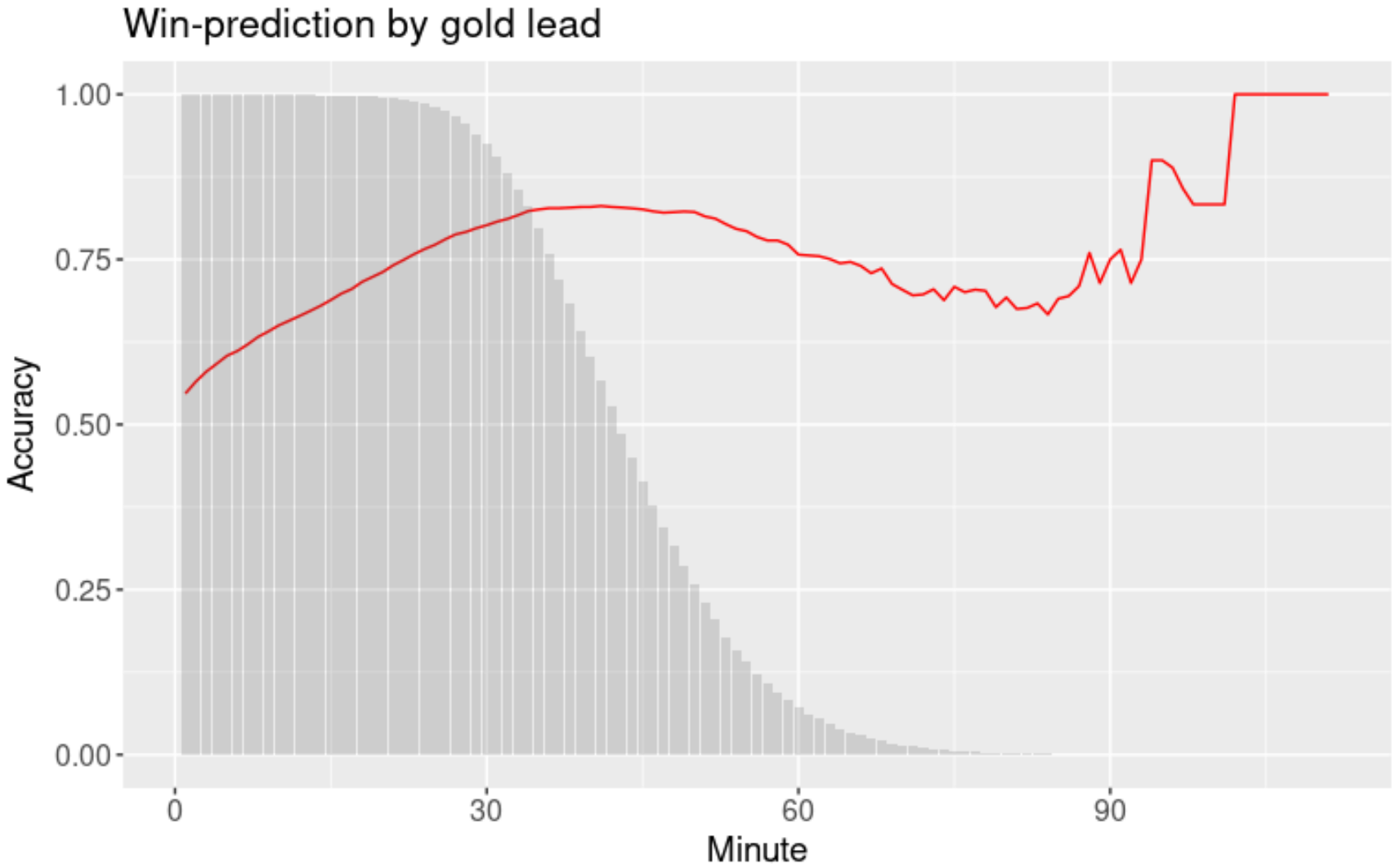
- Looking at a single game (sampled):

	#rows/#columns
- „match.csv“:	1/13
- „players.csv“:	10/73
- „player_time.csv“:	56/32
- „objectives.csv“:	28/9
- „ability_upgrades.csv“:	207/5
- „teamfights.csv“:	13/5
- „teamfights_players.csv“:	130/8
- „purchase_log.csv“:	434/4
- „chat.csv“:	35/5
- Total:	914/~150

Ideas & Challenges

- General win-prediction:
  - Use player history and team composition to predict wins
- On-line win-prediction:
  - Use all available in-game data to predict
  - Update prediction each time new information is available
  - Challenge:
    - Very different types of information
    - Many variables
    - Information is not equispaced
- On-line prediction for events in general:
  - Similar approach as above, while we do not limit us to the binary outcome „win“


Gold per team
5 min
15 min
30 min
60 min
Team 2 Gold
150000
100000
50000
0
0
50000
100000
150000
Team 1 Gold
Team 1 wins
FALSE
TRUE


Win-prediction by gold lead
1.00
0.75
0.50
0.25
0.00
Accuracy
0
30
60
90
Minute

How relevant is the chat?

Winning teams

Losing teams