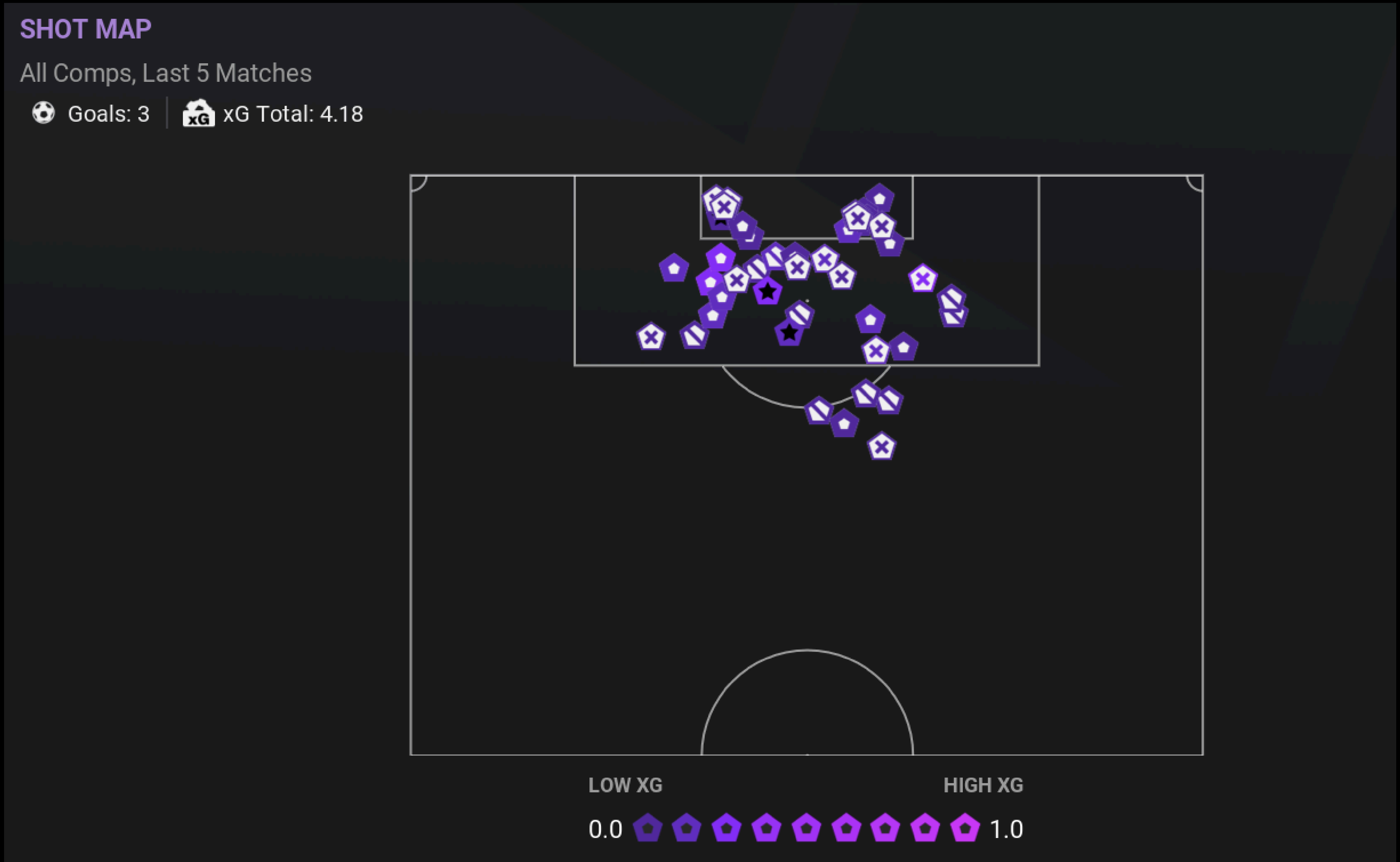
FOOTBALL

THE BEAUTIFUL GAME ORACLE

OBJECTIVE: A DATA-CENTRIC APPROACH TO PREDICT FOOBTBALL MATCHS IN ENGLISH PREMIER LEAGUE (TRACK 2)

SOME KEY FOOTIE TERMS.



xG – Expected goals
xGa – Expected goals against
Shot vs Shot on target

Possible results for a team:
Win/Draw/Loss

Home advantage

Players are considered assets for teams so they have a "Transfer/market value"

---

xG provides a good idea of how many goals a team should score based on the quality of the shots of target. Its a good indication of chances, it doesn't always mean the team will score the amount it stated.

---

passionate football fans wanting to look at the sport from different lenses

- Test the accuracy of bookmakers
- Determine what the best metrics are to predict success in a football team
- Predict better than bookmakers

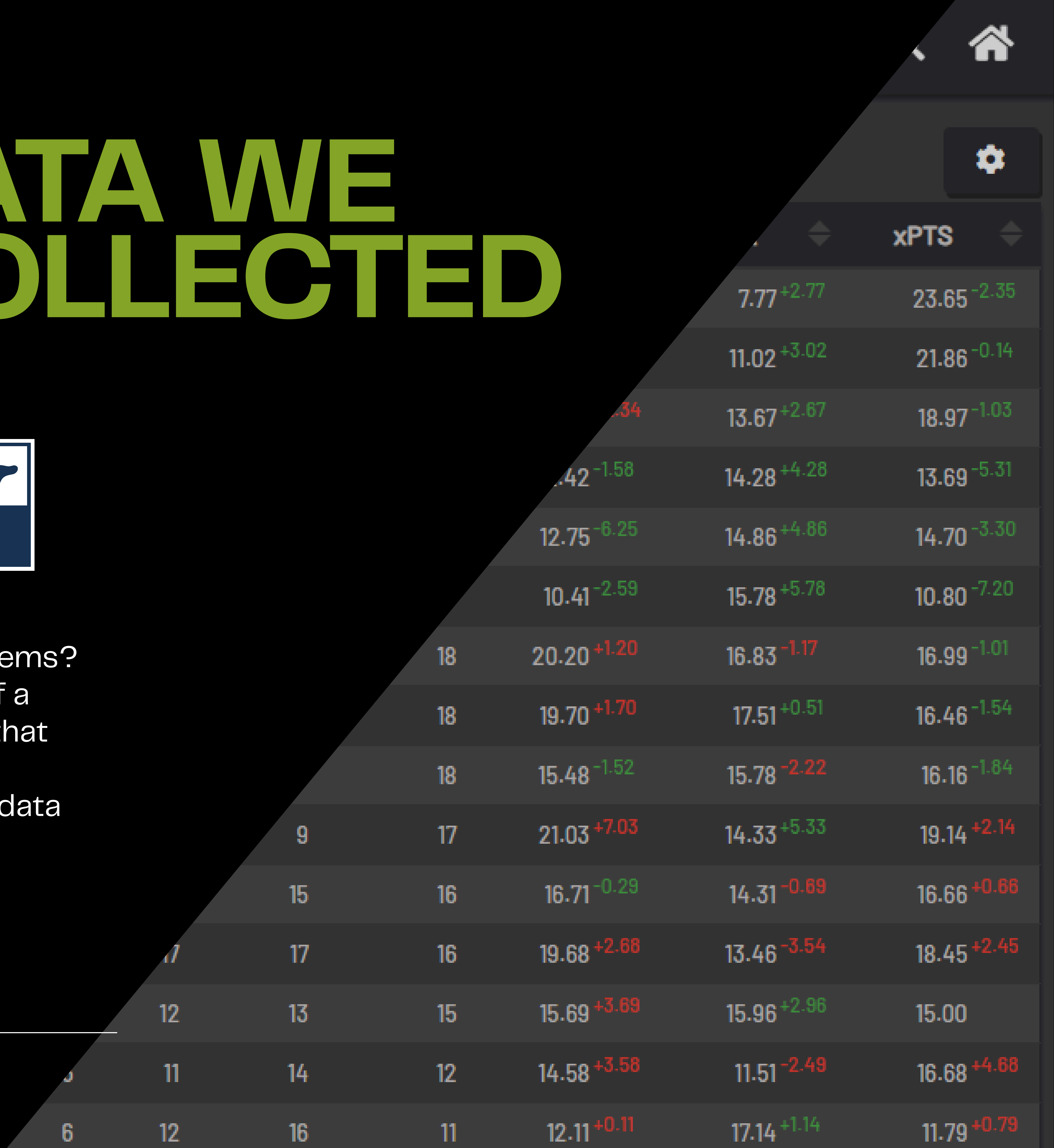
OUR MOTIVATION

ABOUT THE DATA WE COLLECTED

Team match data from the 2022/2023 season to the current season

Each match:
xG, Shots & Goals for both Home and Away teams

Data Cascades problems? just one! The result of a match is something that can't be mistaken.
* except for financial data regarding wages.



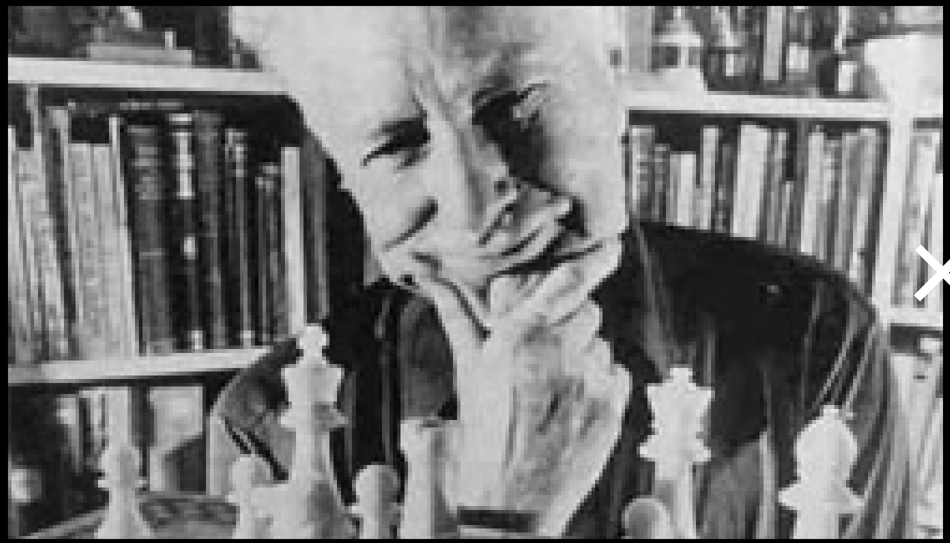
ELO!

ELO!!

ELO!?!

	125	54	20	51	1511.96	-2.12	-9.02	+18.74	-8.93
	125	57	24	44	1509.87	+2.11	+0.19	+12.28	+11.06
	125	36	39	50	1500.63	+11.05	+3.08	-18.74	-7.49
	125	46	26	53	1472.52	-11.06	+11.44	-5.45	-6.84
	125	39	31	55	1465.75	+11.92	-0.19	-8.08	-15.93
	49	10	12	27	1456.44	-11.92	-12.04	+5.78	-12.01
	49	8	10	31	1453.93	-7.69	-10.60	+8.78	+12.01

Elo: A rating system where teams gain/lose points based on results and match metrics



Creator of elo rating system Arpad Elo

Elo is fantastic to evaluate a teams stature across a long period and over many games

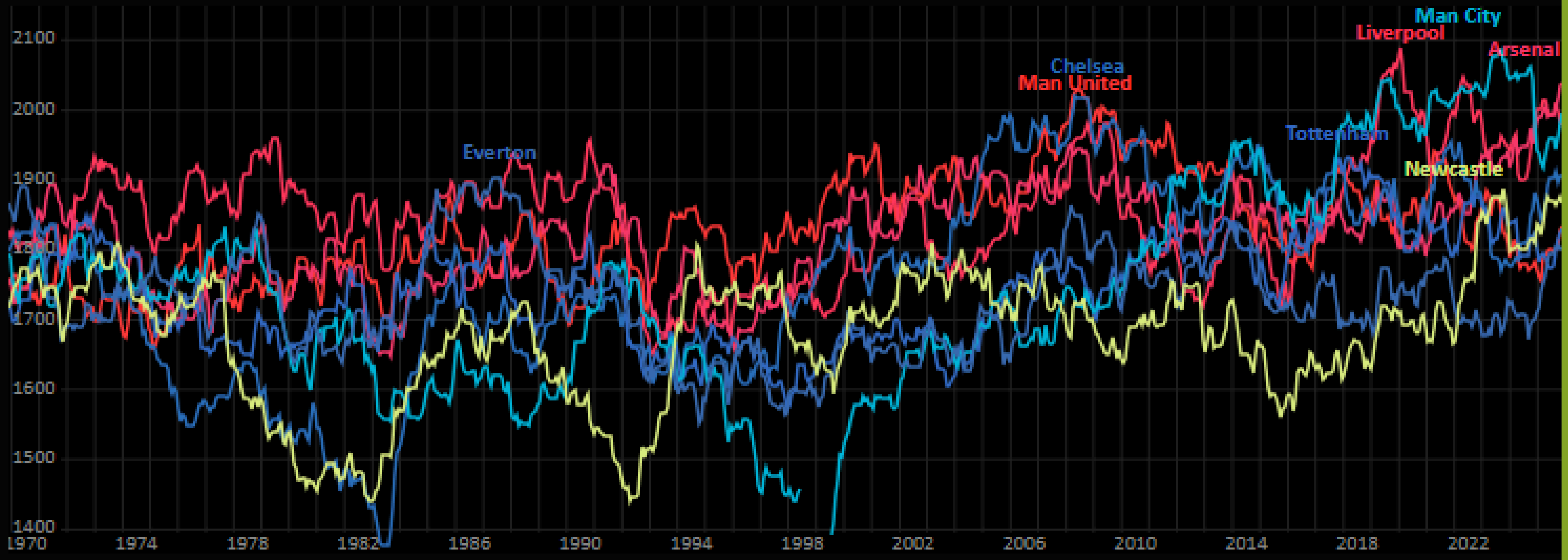
team	played	wins	draws	losses	elo	L5_1	L5_2	L5_3	L5_4	L5_5	next_opponent	next_date	opp_elo	elo_diff
Arsenal	125	82	27	16	1629.99	-1.42	+10.60	+5.03	+6.84	+4.95	Tottenham		1511.96	+118.03
Manchester City	125	84	21	20	1618.07	+13.36	+8.99	-8.46	+7.49	+7.35	Newcastle United		1513.74	+104.33
Liverpool	125	74	29	22	1573.20	-13.36	+9.84	-8.56	-11.06	-7.94	Nottingham Forest		1465.75	+107.45
Chelsea	125	55	31	39	1568.19	+7.04	+9.02	-10.23	+15.94	+7.93	Burnley		1453.93	+114.26
Aston Villa	125	62	27	36	1551.68	+16.50	-9.84	+8.46	+8.93	+4.95	Leeds		1456.44	+95.24
Crystal Palace	125	41	41	43	1543.34	-2.40	+10.45	-5.03	-2.11	-7.53	Wolverhampton Wanderers		1402.97	+140.37
Brighton	125	50	37	38	1533.50	+2.40	+12.03	-12.27	+7.16	-0.70	Brentford		1520.92	+12.58
Brentford	125	46	32	47	1520.92	+12.02	-10.45	+8.56	+13.20	-7.35	Brighton		1533.50	-12.58
Sunderland	11	5	4	2	1520.28	+1.42	-3.08	+10.23	+8.17	-11.83	Fulham		1472.52	+47.76
Bournemouth	125	44	29	52	1517.16	-16.51	-8.99	+8.08	+2.11	+9.16	West Ham		1447.62	+69.54
Newcastle United	125	60	29	36	1513.74	-12.01	-14.71	+5.45	-7.17	+8.55	Manchester City		1618.07	-104.33
Tottenham	125	54	20	51	1511.96	-2.12	-9.02	+18.74	-8.93	+8.66	Arsenal		1629.99	-118.03
Manchester United	125	57	24	44	1509.87	+2.11	+0.19	+12.28	+11.06	+11.84	Everton		1500.63	+9.24
Everton	125	36	39	50	1500.63	+11.05	+3.08	-18.74	-7.49	+7.53	Manchester United		1509.87	-9.24
Fulham	125	46	26	53	1472.52	-11.06	+11.44	-5.45	-6.84	-9.16	Sunderland		1520.28	-47.76
Nottingham Forest	125	39	31	55	1465.75	+11.92	-0.19	-8.08	-15.93	-8.55	Liverpool		1573.20	-107.45
Leeds	49	10	12	27	1456.44	-11.92	-12.04	+5.78	-12.01	-8.65	Aston Villa		1551.68	-95.24
Burnley	49	8	10	31	1453.93	-7.69	-10.60	+8.78	+12.01	-4.95	Chelsea		1568.19	-114.26
West Ham	125	39	28	58	1447.62	+7.68	+14.71	-5.78	-13.20	-4.95	Bournemouth		1517.16	-69.54
Wolverhampton Wanderers	125	36	23	66	1402.97	-7.03	-11.43	-8.78	-8.18	+0.70	Crystal Palace		1543.34	-140.37

team	played	wins	draws	losses	elo
Arsenal	125	82	27	16	1629.99
Manchester City	125	84	21	20	1618.07
Liverpool	125	74	29	22	1573.20
Chelsea	125	55	31	39	1568.19
Aston Villa	125	62	27	36	1551.68
Crystal Palace	125	41	41	43	1543.34
Brighton	125	50	37	38	1533.50
Brentford	125	46	32	47	1520.92
Sunderland	11	5	4	2	1520.28
Bournemouth	125	44	29	52	1517.16
Newcastle United	125	60	29	36	1513.74
Tottenham	125	54	20	51	1511.96
Manchester United	125	57	24	44	1509.87
Everton	125	36	39	50	1500.63
Fulham	125	46	26	53	1472.52
Nottingham Forest	125	39	31	55	1465.75
Leeds	49	10	12	27	1456.44
Burnley	49	8	10	31	1453.93
West Ham	125	39	28	58	1447.62
Wolverhampton Wanderers	125	36	23	66	1402.97

Ours compared To clubelo.com

England

Home Field Advantage: 53.7 Elo points.
Biggest HFA: 130.2 (on Mar 26th, 1983), smallest HFA: 18.6 (on May 16th, 2021).



Ranking

Rank	Club	Elo	Last Match	Coach
1	1 Arsenal	2038	2025-11-08	Mikel Arteta (since 2019-12-22)
2	2 Man City	1997	2025-11-09	Pep Guardiola (since 2016-07-01)
3	4 Liverpool	1979	2025-11-09	Arne Slot (since 2024-06-01)
4	9 Chelsea	1907	2025-11-08	Enzo Maresca (since 2024-07-01)
5	10 Aston Villa	1886	2025-11-09	Unai Emery (since 2022-11-01)
6	12 Newcastle	1867	2025-11-09	Eddie Howe (since 2021-11-08)
7	13 Crystal Palace	1857	2025-11-09	Oliver Glasner (since 2024-02-20)
8	16 Brighton	1847	2025-11-09	Fabian Hürzeler (since 2024-07-01)
9	17 Bournemouth	1834	2025-11-09	Andoni Iraola (since 2023-07-01)
10	18 Man United	1834	2025-11-08	Rúben Amorim (since 2024-11-11)
11	19 Tottenham	1830	2025-11-08	Thomas Frank (since 2025-07-01)
12	20 Brentford	1829	2025-11-09	Keith Andrews (since 2025-07-01)
13	23 Everton	1804	2025-11-08	David Moyes (since 2025-01-11)
14	30 Fulham	1779	2025-11-08	Marco Silva (since 2021-07-01)
15	33 Forest	1764	2025-11-09	Nuno Espírito Santo (since 2023-12-20)
16	42 West Ham	1738	2025-11-08	Graham Potter (since 2025-01-09)

ORACLE DECISIONS

Year	Winner	Runners-up
2024-25	**Liverpool**	**Arsenal**
2023-24	Manchester City	Arsenal
2022-23	Manchester City	Arsenal

Data only consists of the previous 3 seasons, ensuring we have sufficient match data, but also have data that may be relevant to the current teams (roughly)

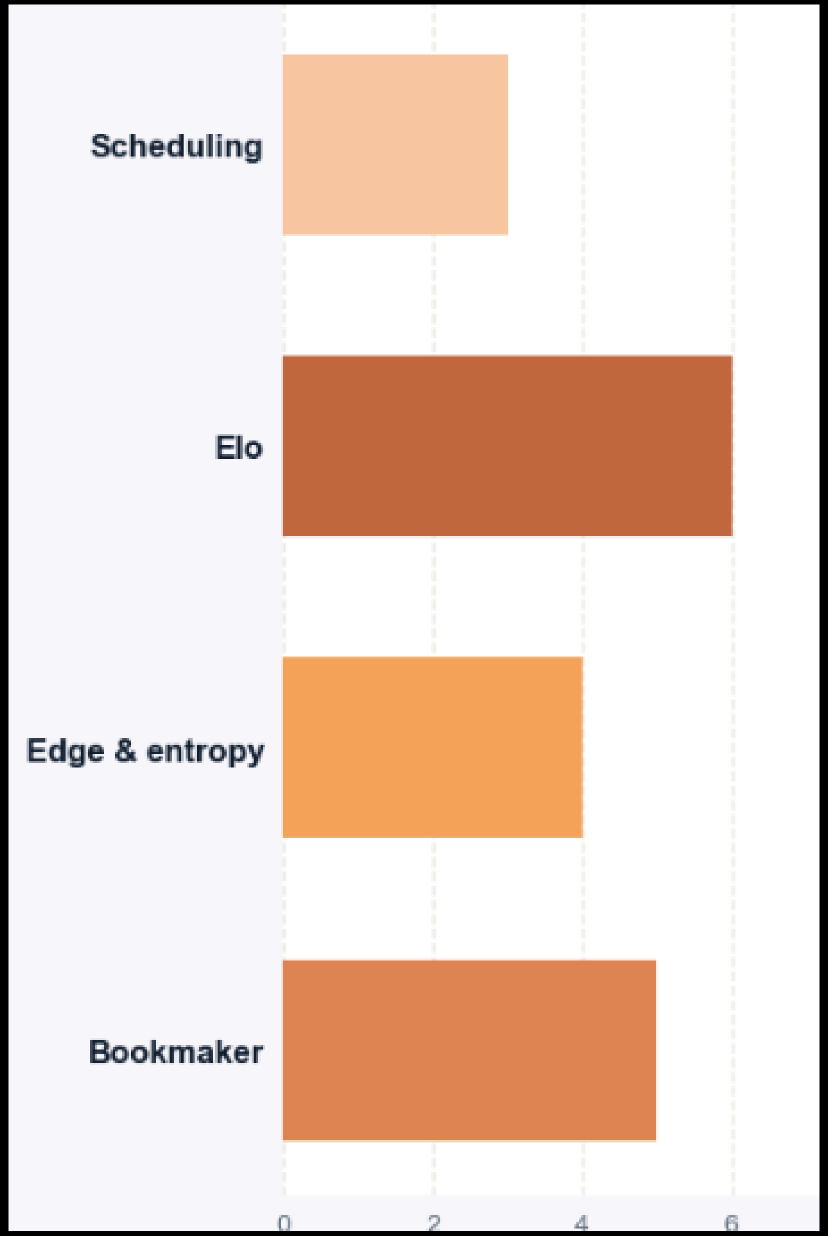
We decided not to use individual player data due to scope of project time frame

Validation data is strictly the 2025/2026 season (Current season) as that is the matches we actually want to predict

IDEAS → LENSES

×

Market View

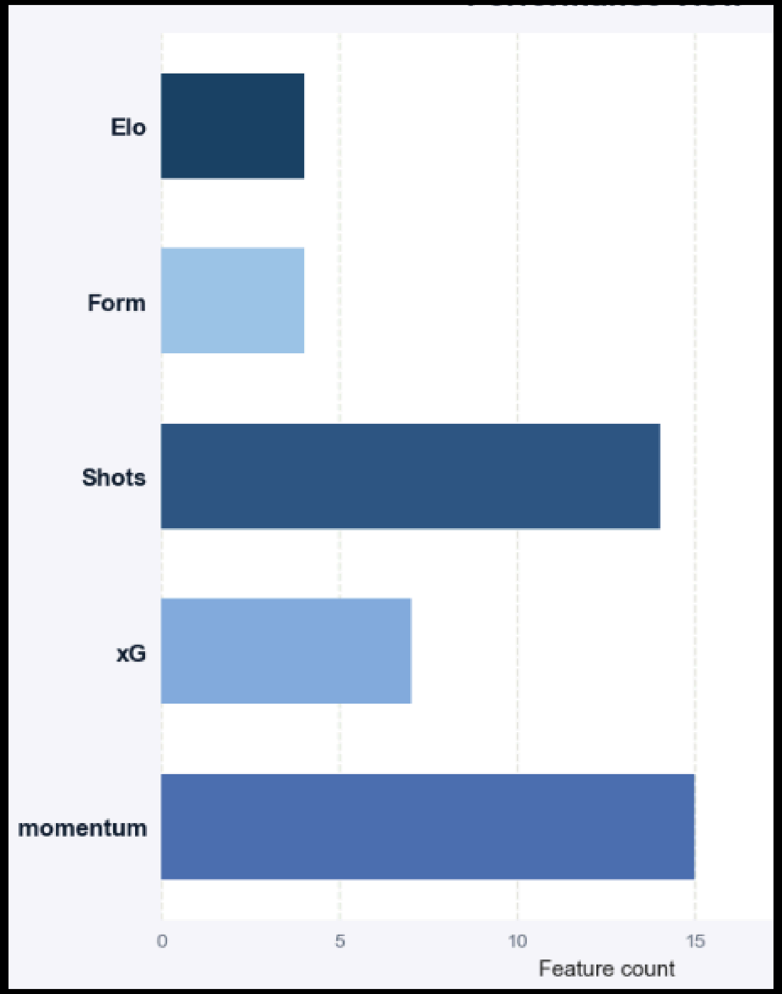


Data is everything to our project, from scraping, formatting, cleaning to feature views and dataset versioning. DataPerf served as our ideological role-model.

3 different data views (Performance Lens, Market Lens, Financial Lens),
3 different feature sets. Each dataset version produces features processed in accordance with a new idea, like Elo, Momentum, Form Volatility, Form Decay, etc.

Models stay the same!

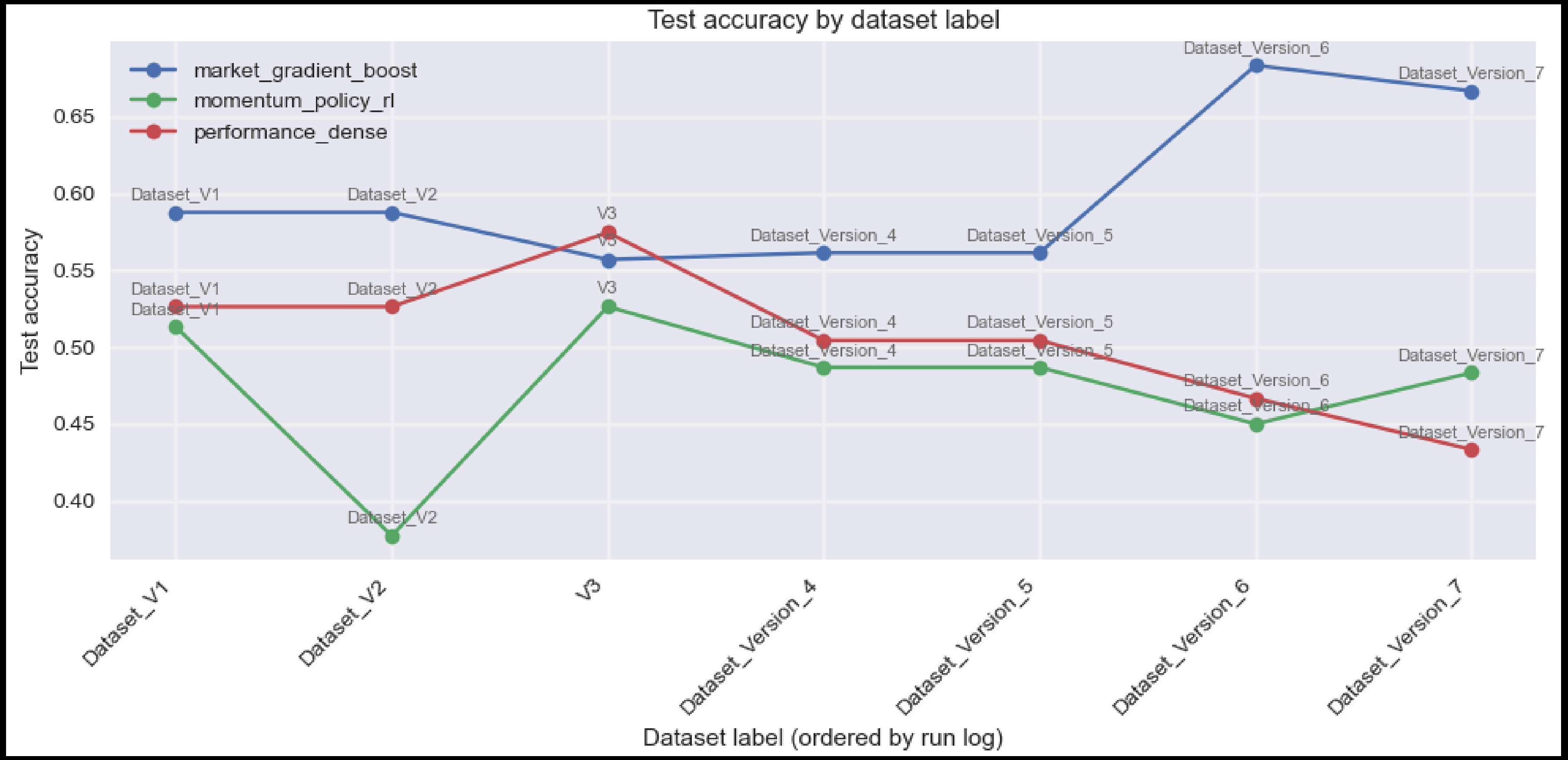
Performance View



DATASET VS DATASET

Methods like rolling five-aggregates, momentum z-scores, Elo, Season gap z-scores, rolling std-dev, form volatility markers, SHAP, LOO and many more!

Went through 7 iterations of datasets with varying techniques and hypotheses being tested.



A TRIP THROUGH DATASETS

V1 → **V2** → **V3** → **V4** → **V5** → **V6** → **V7**

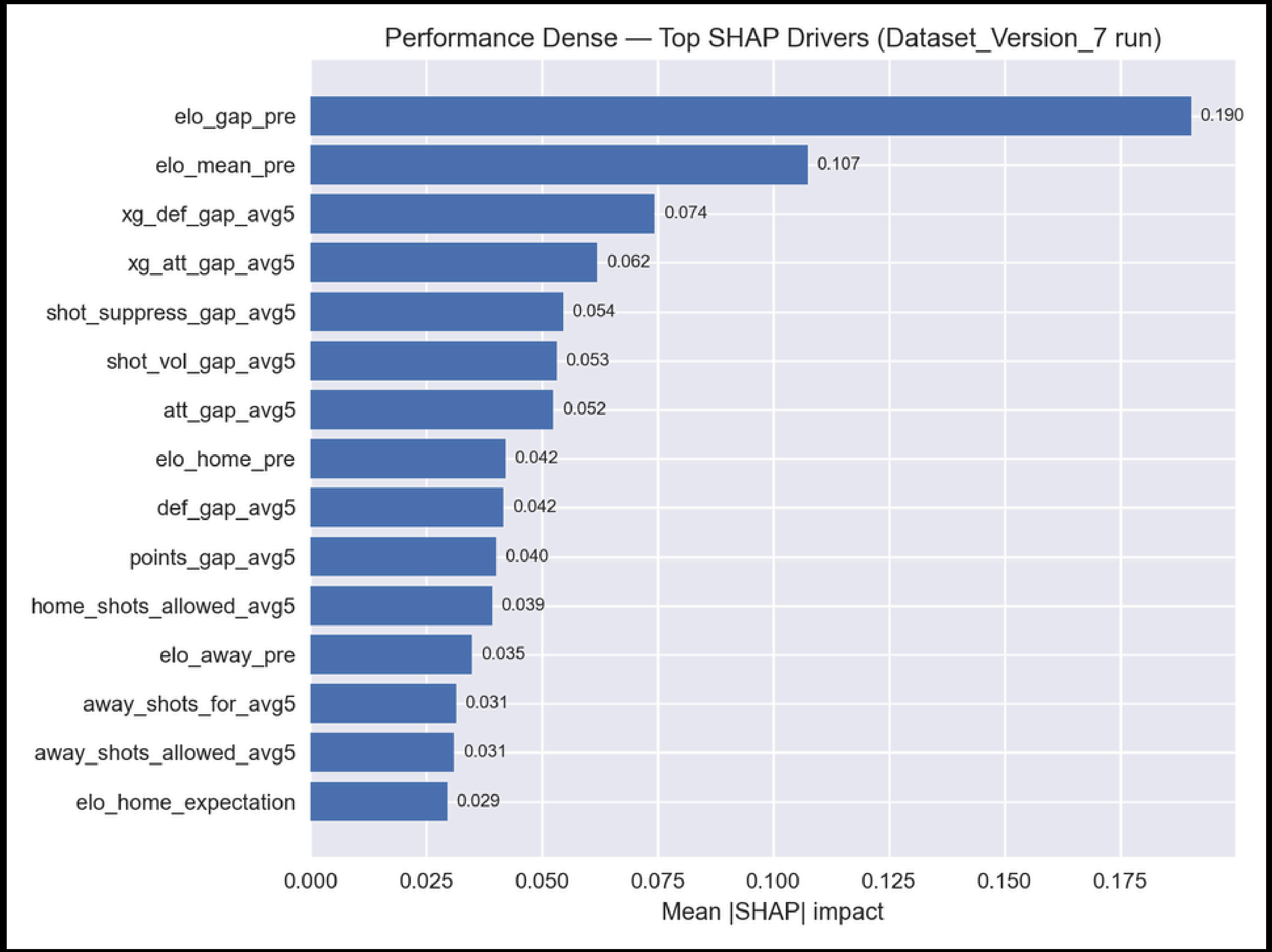
- **V1**: Understat data dumps to neat, flat, consistent data
- **V2**: Adding model-friendly rolling windows, intesity markers and momentum trackers
- **V3**: Calculated Elo and shots data added for context
- **V4**: Updated Elo for patterns, such as preseason effects, recency decay and noise
- **V5**: Comparsion using SHAP and LOO for impact of Elo vs Market Odds
- **V6**: Dynamic train/test split on recent data, Data Pipelines for CI/CD
- **V7**: Form Volatility and Deviation features, Decay of form to smooth predictions.

SHAP STORIES: PERFORMANCE

TLDR: Elo FTW!

ELO gap had the strongest mean SHAP value at 0.19, beating out XG gaps and shot differentials, demonstrating high-value data.

LOO shows accuracy dropping 3.3 points, and revealed features like elo_mean_pre which were negatively impacting accuracy.



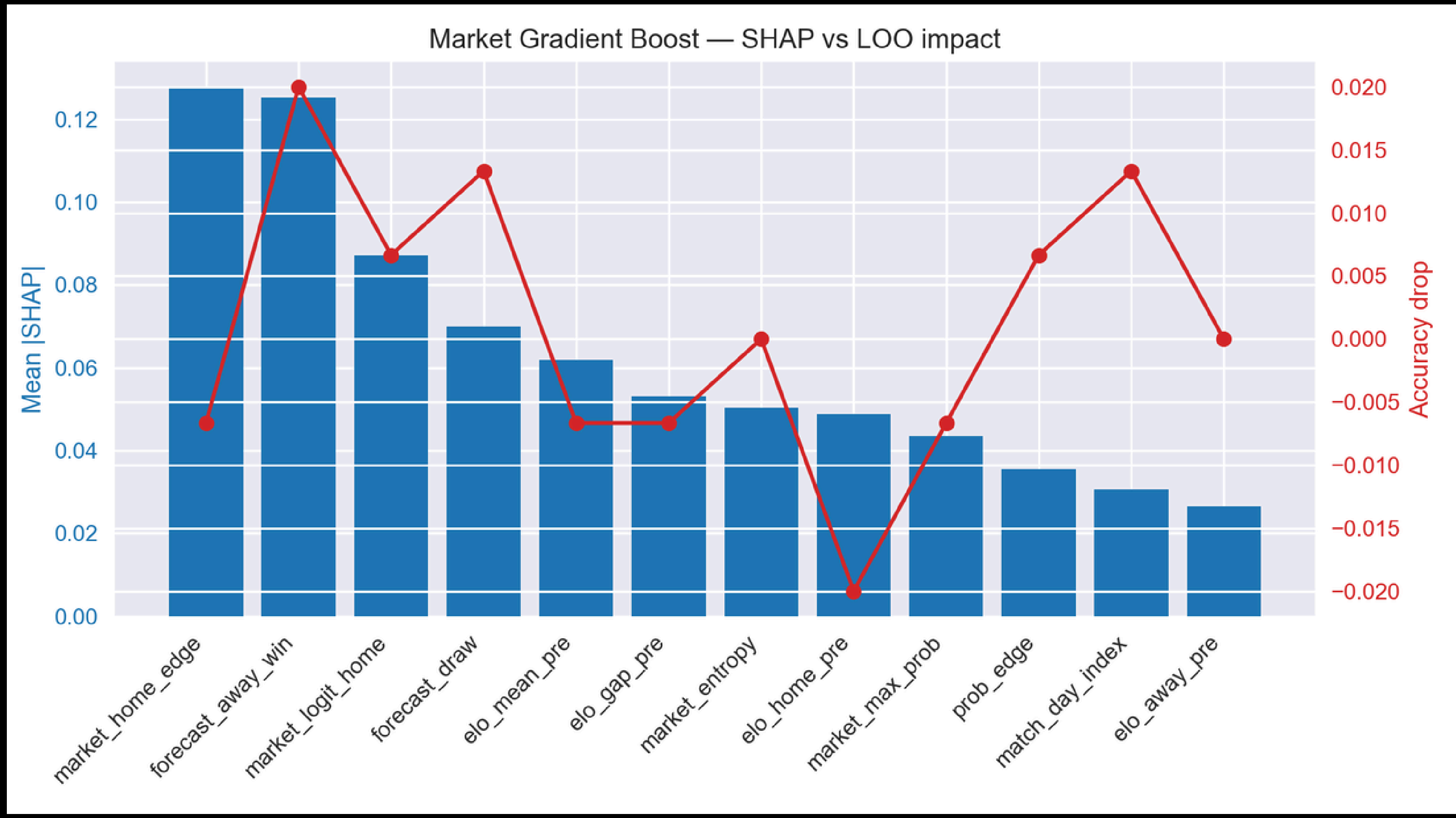
SHAP STORIES: MARKET

×

Still Elo??

Bookmaker odds still dominate as expected, but elo still makes a significant contribution.

Match_day_index has low SHAP but high accuracy drop, indicating a calibration signal rather than a contributor, balances out other features for drift.



We started with the stuff we know happened, and figured out the stuff to happen.

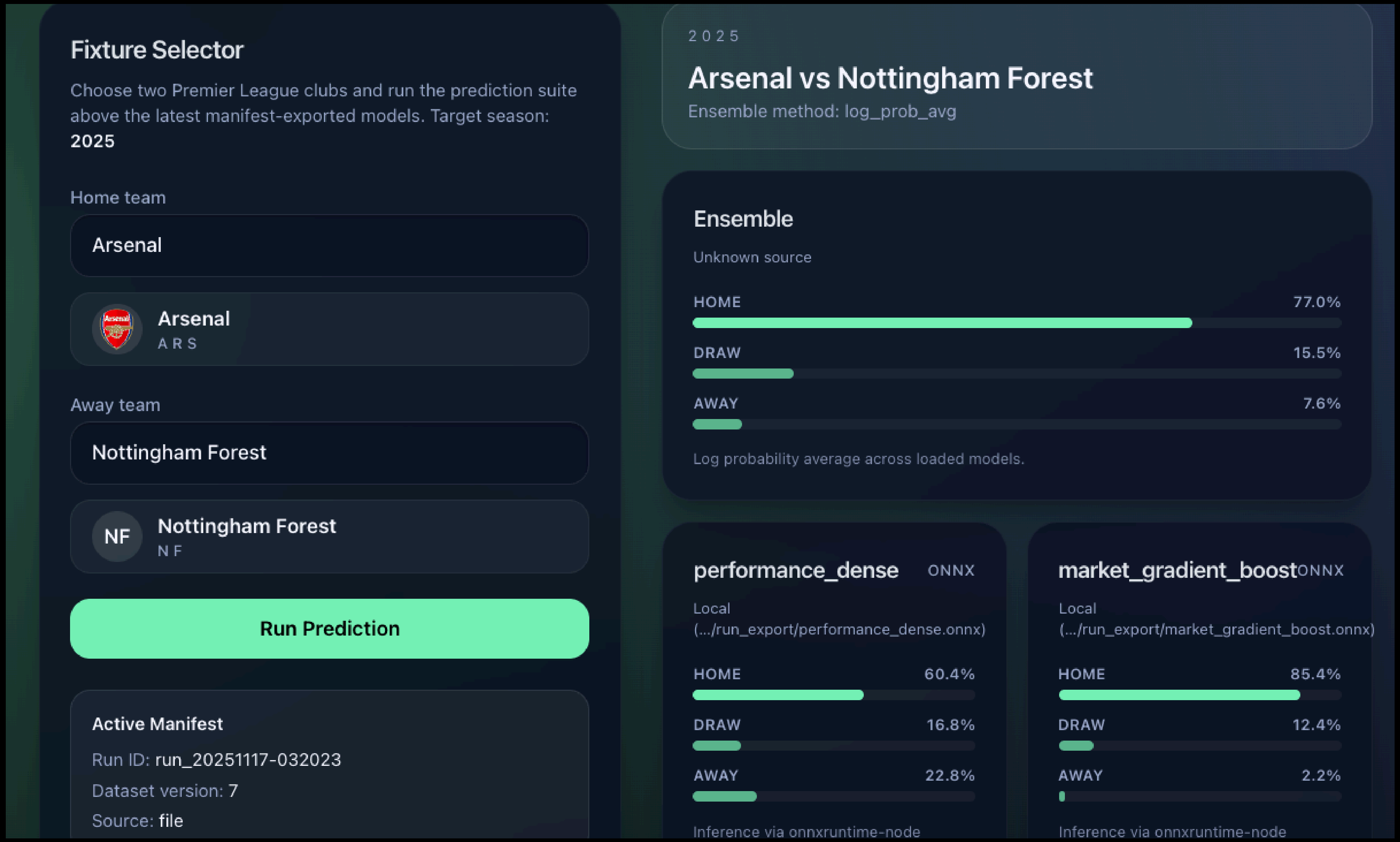
Trained on 22/23 – 24/25 seasons
Tested on current 25/26 season

Predicts on the future

×

Train/Test/Validation Splits remained exactly the same for each Dataset Version to ensure consistent comparison

Caveat: Not all the Dataset Versions have the same split, mostly to prevent undefined leakage from the test set before safeguards.



THE PREDICTION MAGIC

Financial Lens
Capturing Structural Inequality

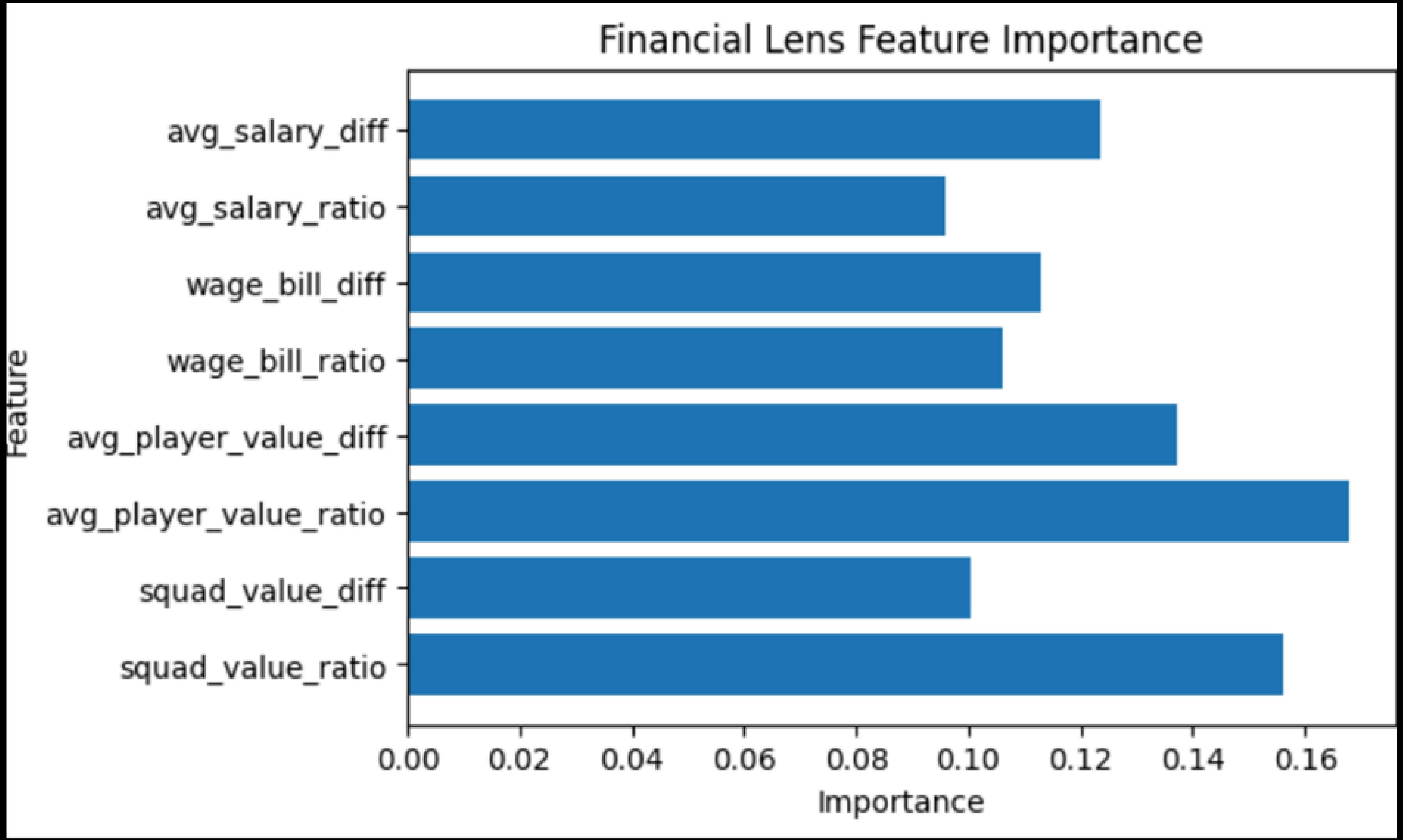
- Built using FBRef wages and Transfermarkt player valuation
- Engineered ratios & differences: squad value gap, salary gap, wage bill ratio
- Reflects long-term structural advantage, not match-specific randomness
- Weak early season → improves as season progresses
- Best at capturing resource gaps, squad depth, and power dynamics

FINANCIAL LENS

FINANCIAL MODEL FEATURES

Financial Lens
Capturing Structural Inequality

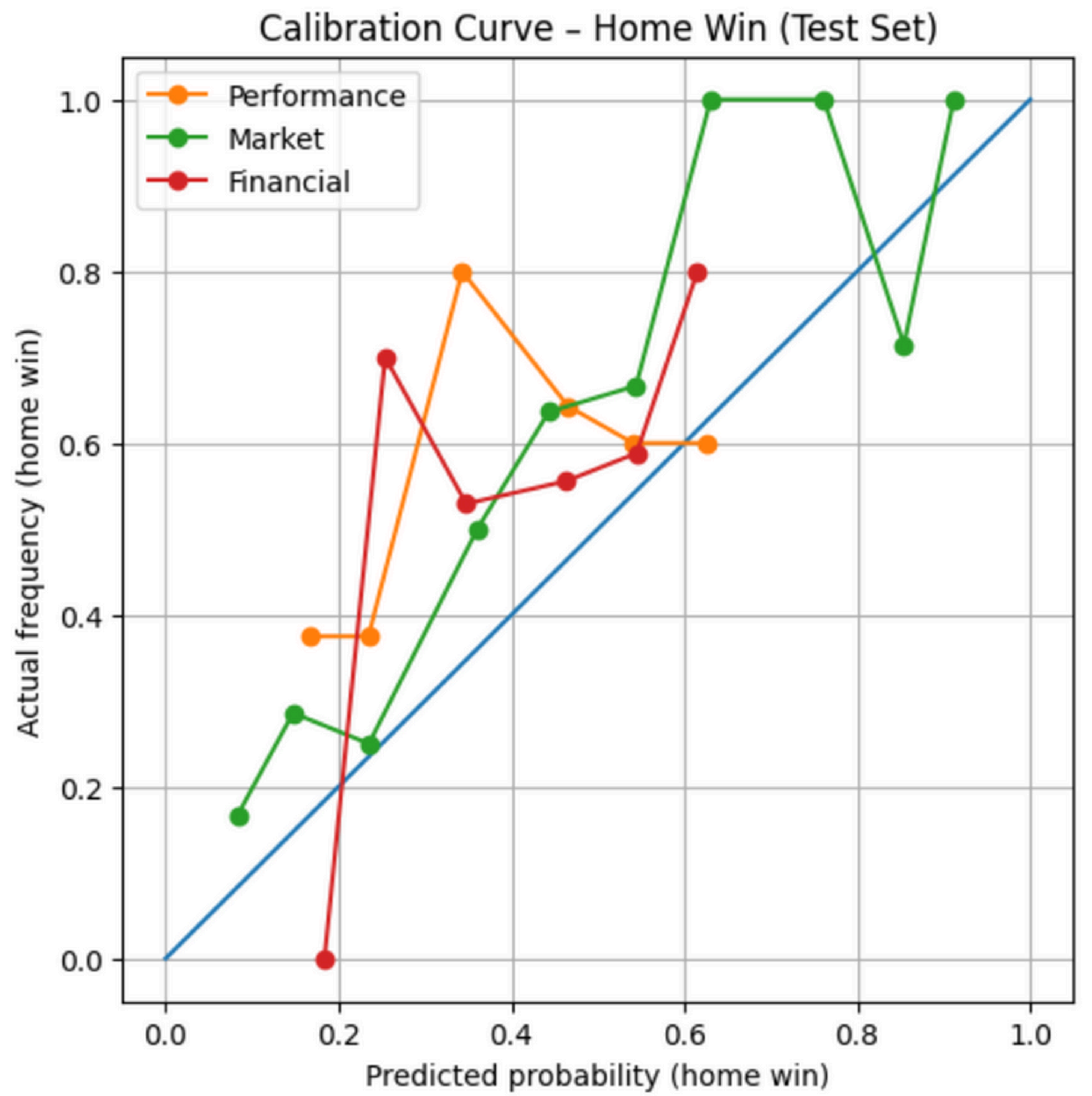
- Squad value ratio = strongest predictor. This makes sense because squad value captures quality, depth, and market perception simultaneously.
- The next strongest signals are average player value difference and ratio. These features identify whether a team has a consistent baseline of talent across its squad.
- Features like Wage bill differences and average salary also matter, but to a lesser extent. Salaries can be noisy since some teams overpay, some underpay, and contract timing affects numbers, so these features are weaker.



CALIBRATION CURVE

How Honest Are the Probabilities?

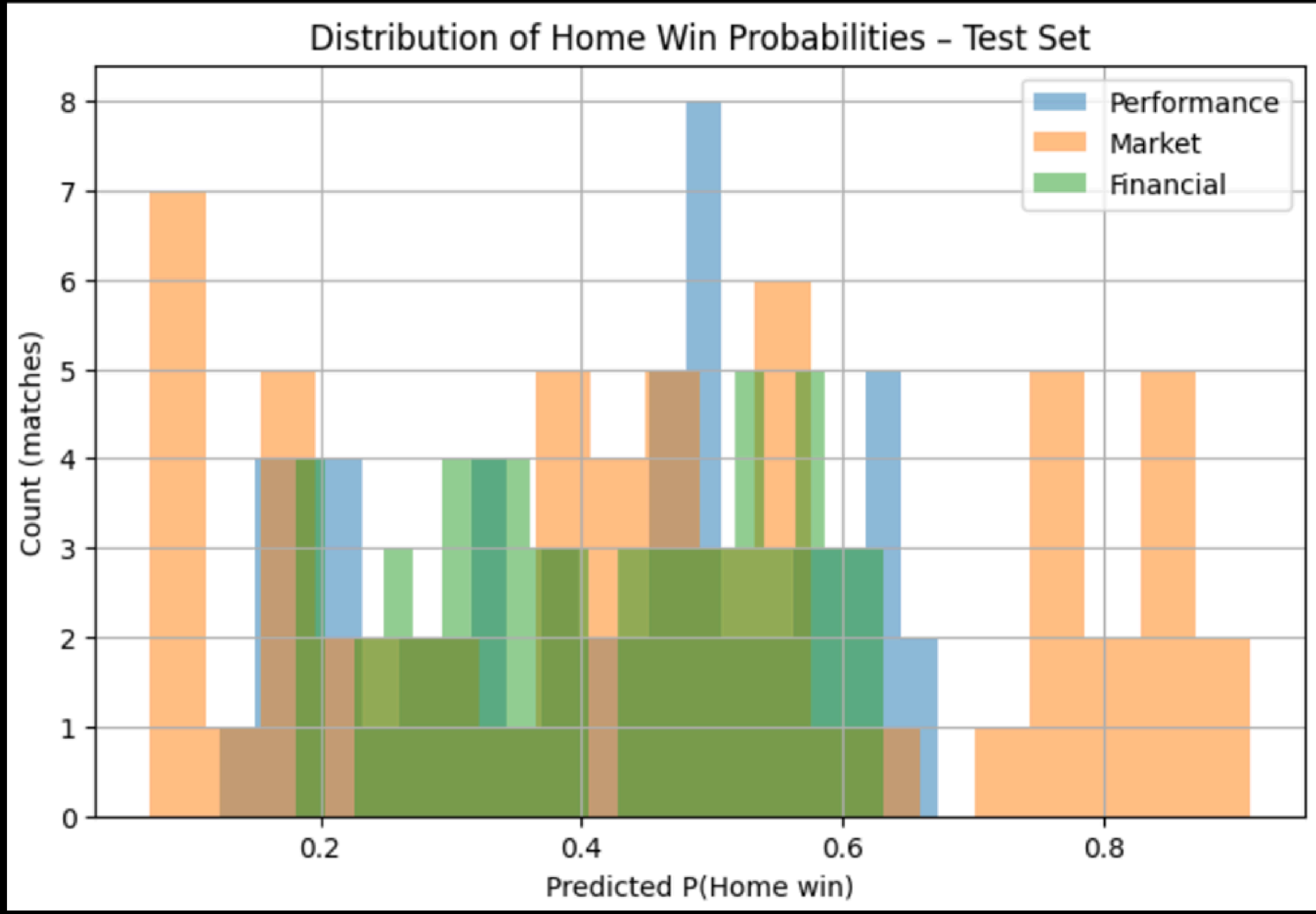
- If a model states that this event occurs 60 percent of the time, calibration checks whether it actually occurs at that rate.
- Market Lens is the best calibrated across the probability bins.
- Performance lens fluctuates → reacts to noisy match events.
- The financial lens is underconfident early, but improves in the mid-range.
- Calibration reveals model trustworthiness, not just accuracy
- Overconfidence vs underconfidence reflects each lens' worldview



PROBABILITY DISTRIBUTION

How Each Lens Expresses Confidence (Home Win Probabilities)

- The market lens is the sharpest. It often predicts very high or very low probabilities.
- The performance lens tends to stay in the middle. It rarely goes extreme because match events are noisy.
- The financial lens is the most conservative. Its probabilities are tightly centred because structural inequality doesn't guarantee individual match outcomes.
- Market thinks in bold strokes, performance thinks in recent events, and financial thinks in long-term power.

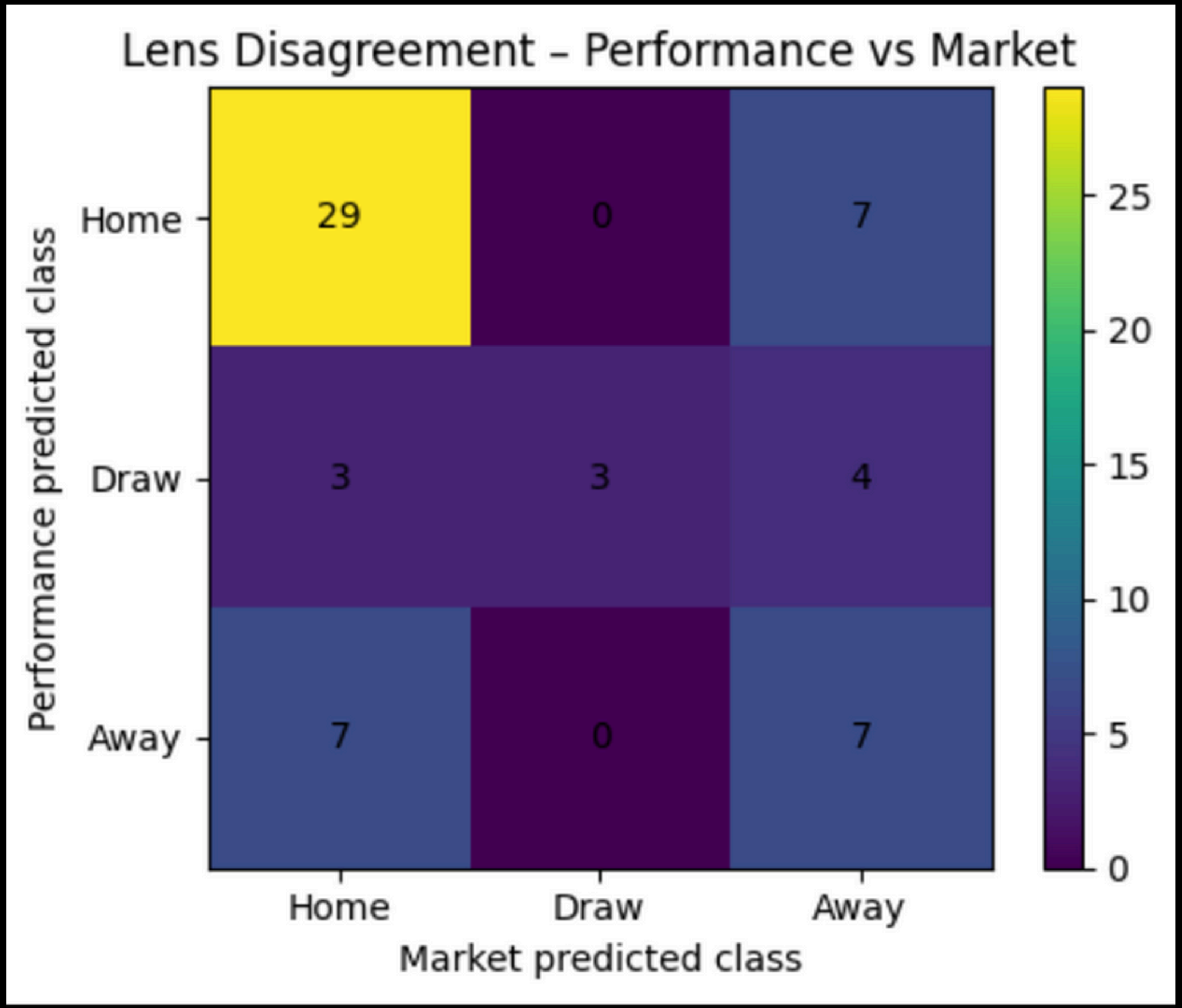


LENS DISAGREEMENT

Where the Lenses Disagree (Performance vs Market)

- Rows = Performance predictions and Columns = Market predictions
    - e.g. (row = Home, column = Home) = both predicted home
    - (row = Home, column = Away) = performance predicted home but market predicted away

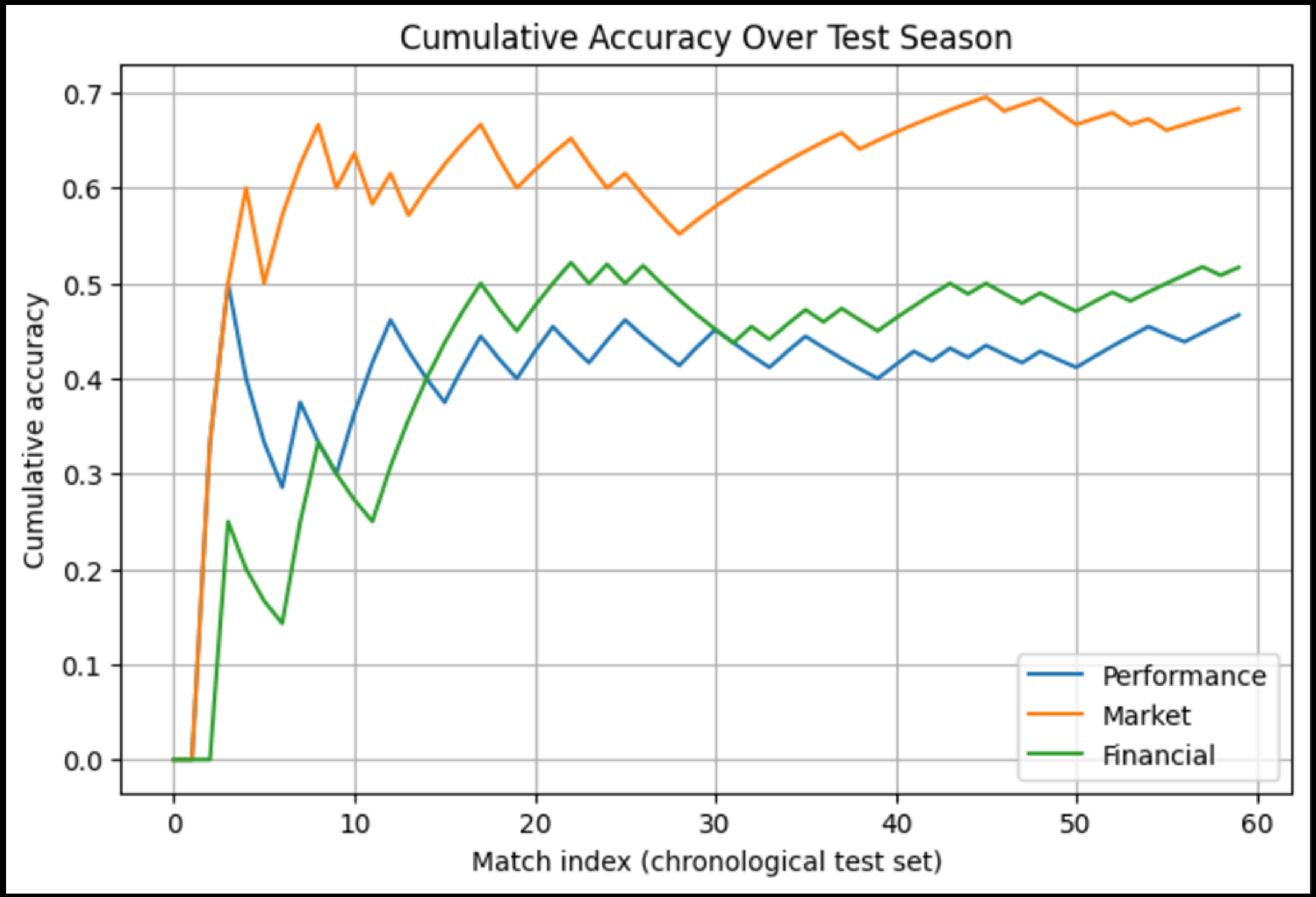
- Major disagreement on draws and close matches
- Performance predicts draws more frequently
- Market avoids mid-range outcomes → prefers extremes
- This disagreement is one of the most significant outcomes of our project. It shows why separating lenses matters. By treating datasets as unique perspectives, we can see where human judgment, match events, and financial power diverge



CUMULATIVE ACCURACY OVER TIME

How Accuracy Evolves Across the Season

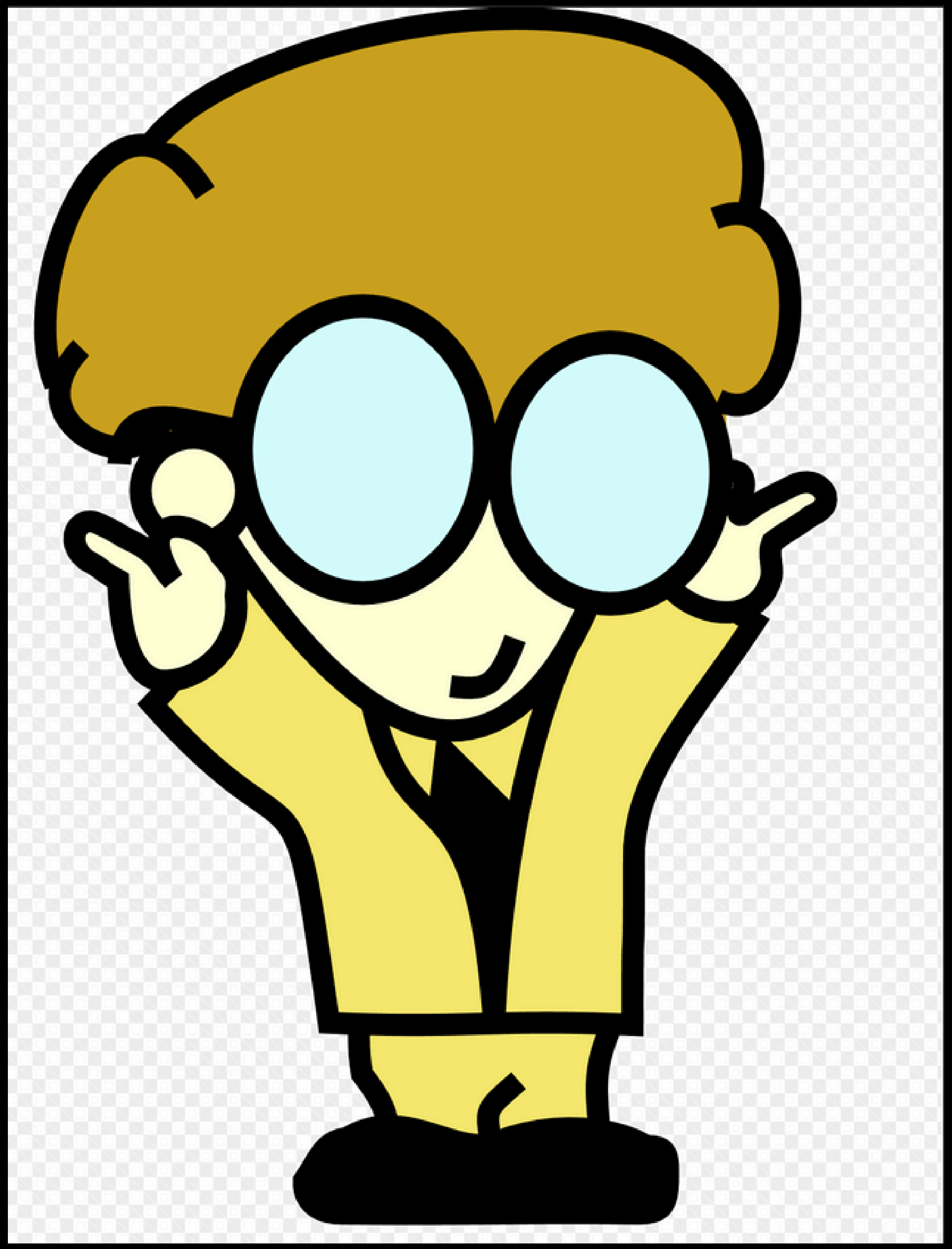
- The market lens is the most stable. It starts the season strong and stays strong
- The performance lens is much more volatile.
- The financial lens is the most interesting. It starts out weak because financial strength doesn’t automatically show up in week one. However, as the season progresses, the advantages of squad depth, wage bills, and valuations become increasingly important. That’s why its accuracy climbs steadily over time.
- Performance reflects short-term events, markets reflect immediate expert judgment, and financial data reflects long-term structural power.”



KEY INSIGHTS

Key Insights Across All Lenses

- The first insight is that data choice mattered more than model choice. Even though all three lenses used similar model architectures, they behaved completely differently because the underlying data encoded different assumptions about what matters in football.
- The market lens turned out to be the most calibrated and the most stable across the season
- The performance lens captures real-time on-pitch actions, allowing it to identify momentum and event-driven signals. But because match events are noisy, this lens is also the most volatile.
- The financial lens is weak at the start of the season but gets better over time. slow-burn effect
- lens disagreement is meaningful. When performance predicts draws and market predicts home wins, it tells us something about how different forms of knowledge interpret the same match. This connects back to course themes about multiple ways of knowing, pluralism, and understanding the assumptions behind data.



THANK YOU

www.reallygreatsite.com