StatsBomb

Project Kickoff Expected Goals (XG)

Mohamad Traiki

StatsBomb

Contents

1. Building an Expected Goals (xG) Model.
2. Understand shot quality.
3. Incorporate defender positioning (freeze-frames).
4. Test generalisation across tournaments.
5. Compare ML approaches: LR, XGB, scaled variant.
6. Hyper-Parameter Tuning.

StatsBomb

Data-Sets

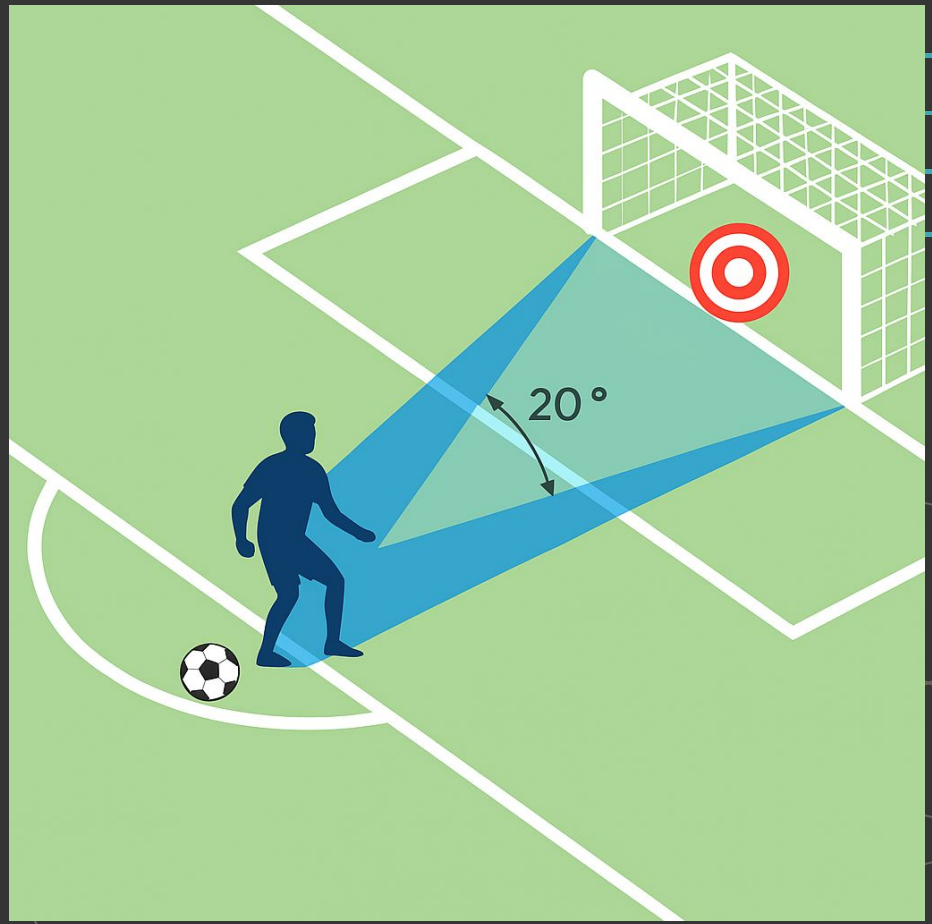
- WC18: ~1,660 shots
- EURO 2020: ~1,250 shots
- WC22: ~1,450 shots

Features

- Shot type / technique / body part
- Location (x, y)
- Distance, angle
- Defenders_between
- Min_defender_dist
- Defenders_in_cone

Target

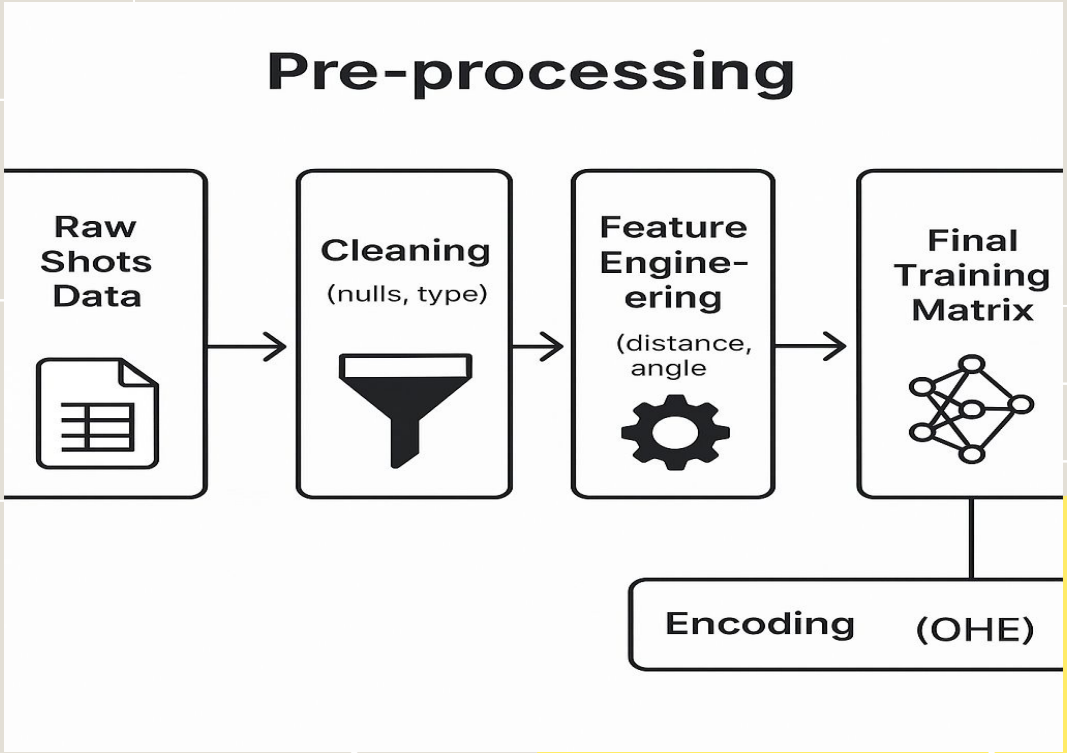
- is_goal



StatsBomb

Preprocessing

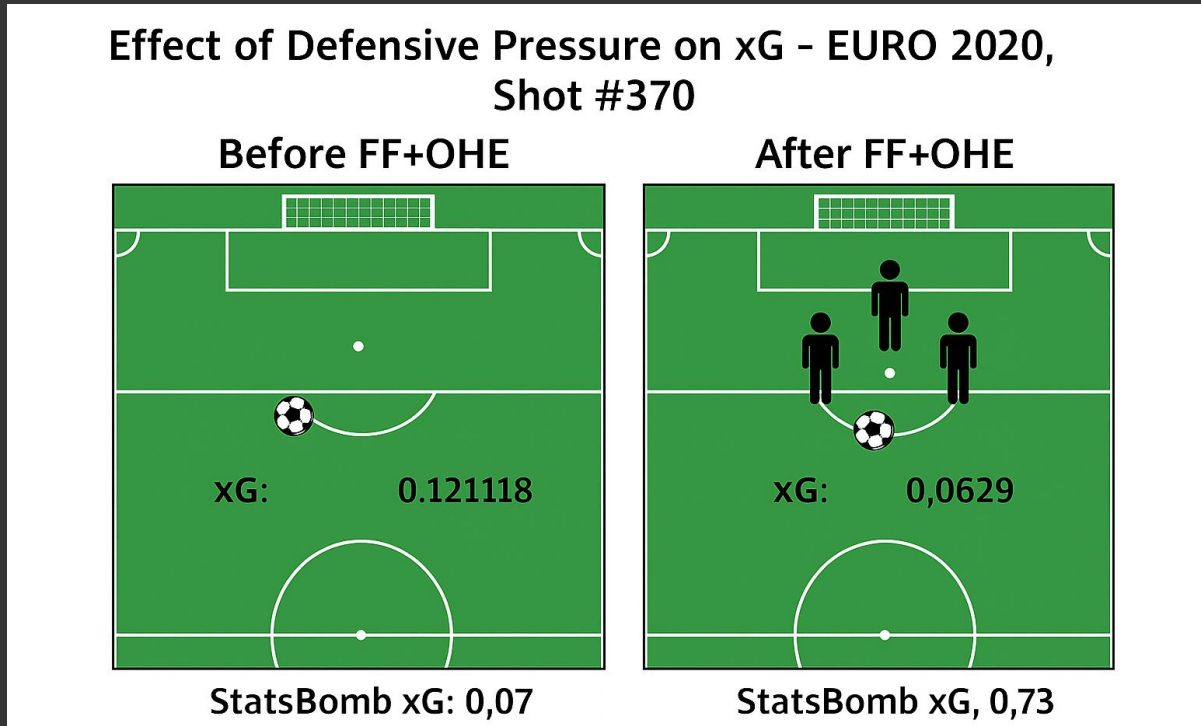
- Removed shootout penalties
- Extracted location → (x, y)
- Engineered geometric features
- Engineered defensive FF features
- One-Hot Encoding of categorical features
- StandardScaler only for LR_scaled variant



StatsBomb

Freeze-Frame (FF)

- FF captures defensive context
- Model learns realistic difficulty
- More calibrated to real scoring likelihood
- xG becomes more situationally aware
- Better ranking + better calibration



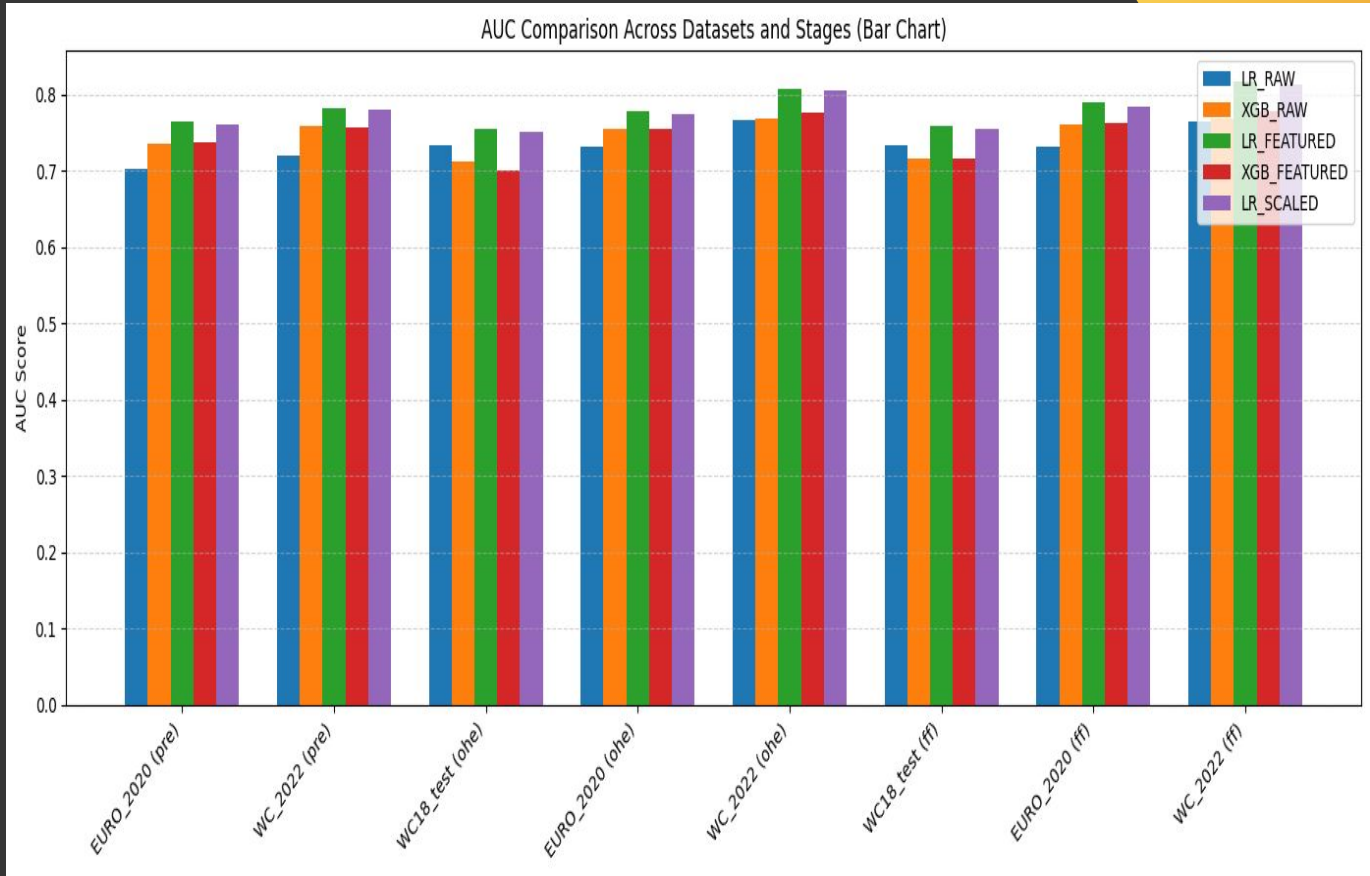
StatsBomb

Models

Baseline:
- LR_raw
- XGB_raw

Enhanced:
- LR_featured
- XGB_featured
- LR_scaled

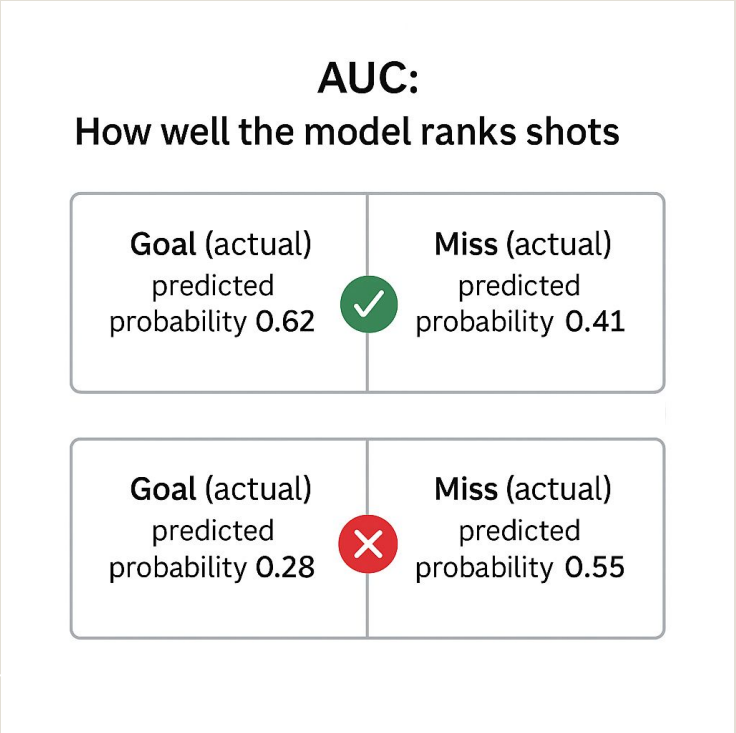
Hyperparameter tuning:
- GridSearchCV
- Best model: LR_tuned



StatsBomb

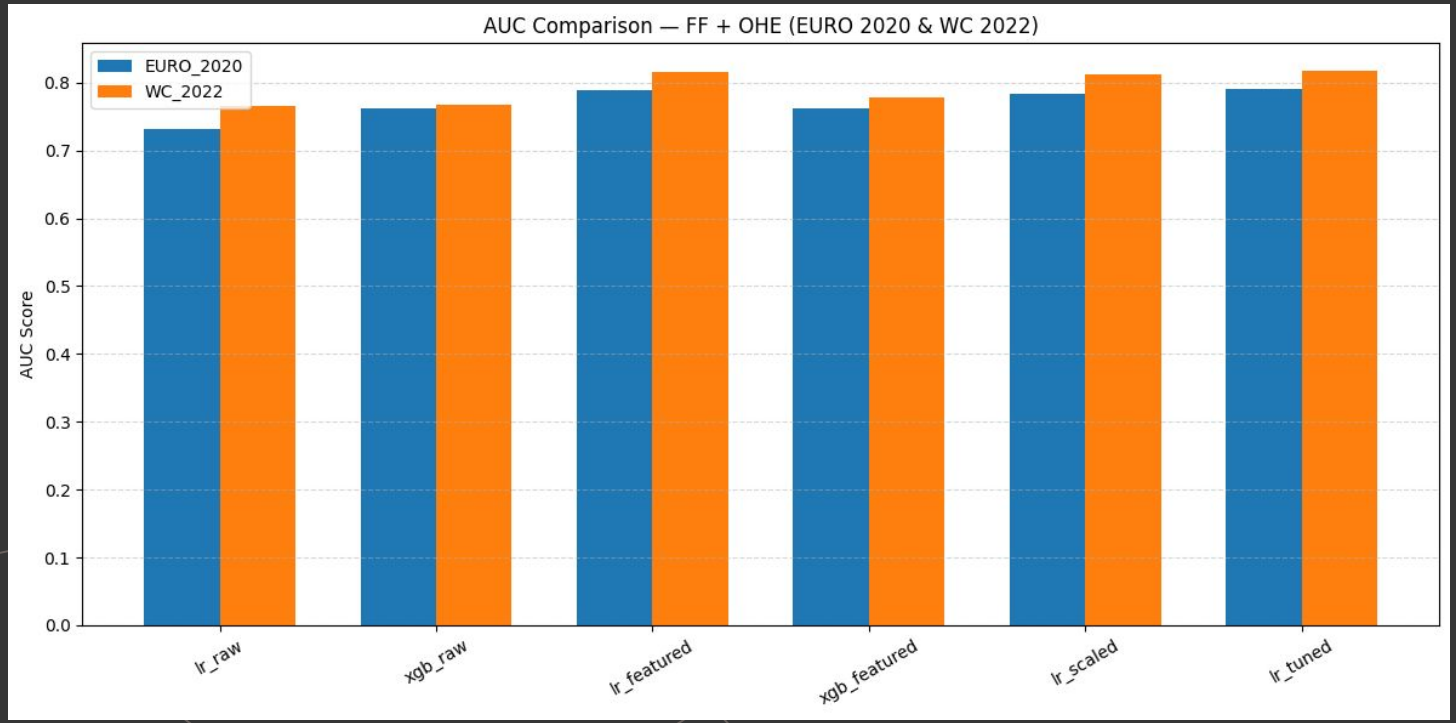
Evaluation Metrics

- **AUC (Area Under ROC Curve)**
  - **Measures how well the model separates goals from non-goals**
  - **Higher = better discrimination ability**
  - **“How good is the model at ranking chances?”**
- **Brier Score**
  - **Measures how accurate the predicted probabilities are**
  - **Lower = better calibration**
  - **“How close are predicted probabilities to reality?”**
- **Total xG Sum (Predicted xG)**
  - **Sum of predicted goal probabilities over all shots**
- **StatsBomb xG (Reference Model)**
  - **Industry benchmark for expected goals**
  - **Used for comparison, not for training**

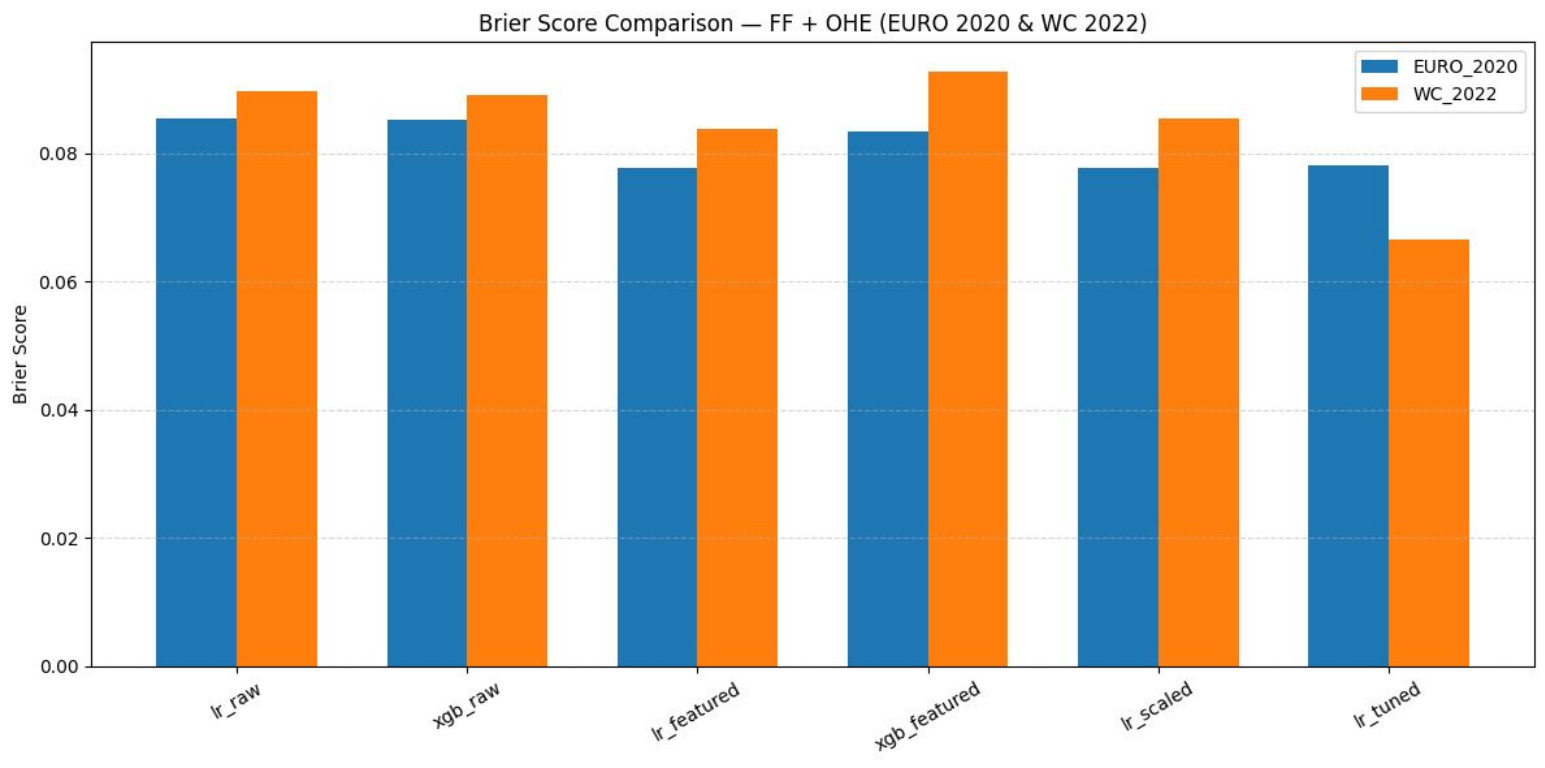


StatsBomb

- # AUC (Area Under ROC Curve)

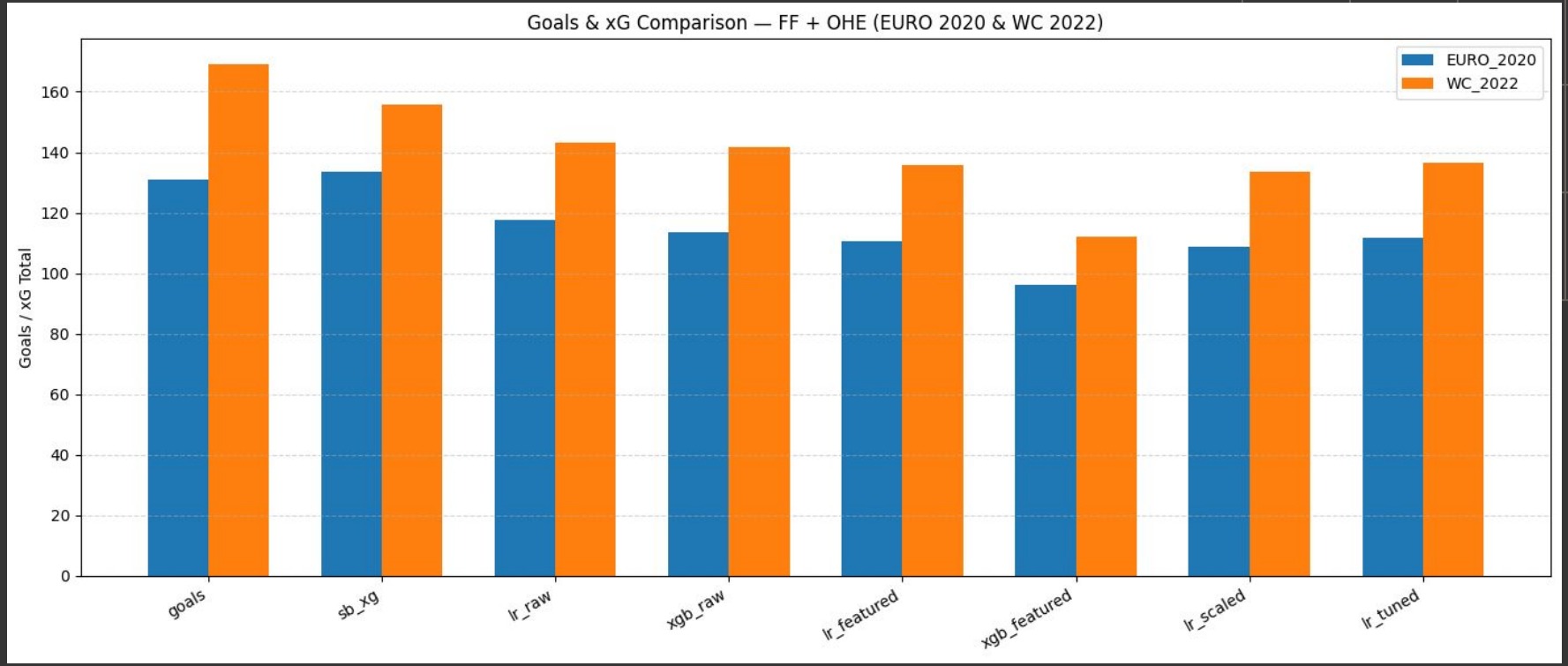


Brier Score


Brier Score Comparison — FF + OHE (EURO 2020 & WC 2022)
EURO_2020
WC_2022
Brier Score
0.00
0.02
0.04
0.06
0.08
lr_raw
xgb_raw
lr_featured
xgb_featured
lr_scaled
lr_tuned

StatsBomb

Goals vs XG



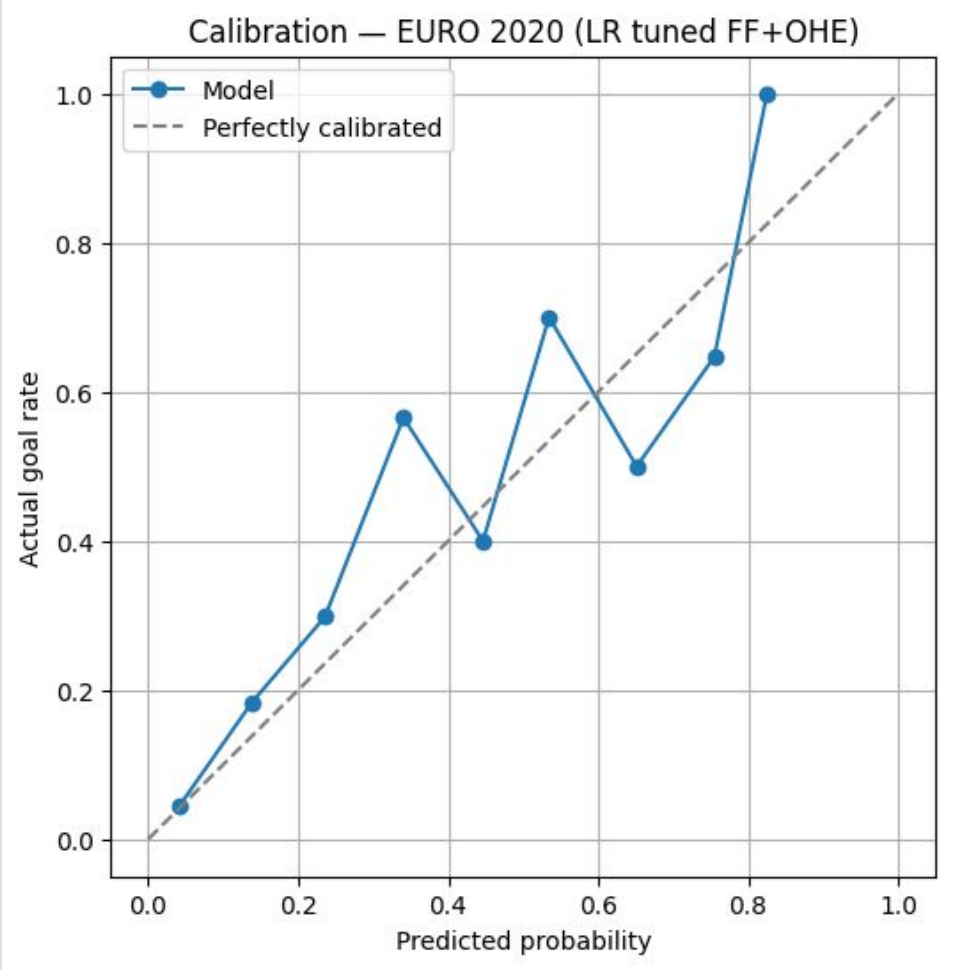
Error Analysis

Findings

- Ranking mostly correct
- Mid-range badly underestimated
- Some bins inconsistent

Possible Reasons

- Defender impact over-weighted
- Good locations under-valued
- Limited training data



StatsBomb

Conclusion

Possible enhancements:

- Train on a much larger dataset to improve model stability and calibration.
- Rebalance or rescale defender-related features to avoid overly penalizing pressure.
- Refine feature engineering for mid-range shots where the model underestimates risk.

Future Considerations:

- Incorporate richer freeze-frame geometry (angles, relative positions, dynamic spacing).
- Experiment with more flexible models such as gradient boosting or neural networks.
- Evaluate model fairness and consistency across different tournament styles and teams.

Thank you