Football Match Outcome Prediction

using DeepSet Player Aggregation

Phase 2

Group 12
Mohammad Bahrami
Sayed Masoud Mousavi

Recap

- Task: Football Match Outcome prediction

- Data: Proprietary Data from Mr. Tarak
  - Contains team and player features per match
  - Needed Cleaning
  - Needed Aggregation for post game features
  - The data is saved after the pre-processing phase and is loaded again (in a ready-to-use state) in this phase.

Outline

- Setup
  - Cross-validation
  - Class Balancing
  - Baselines
- Modeling
- Evaluation

Setup

❖ Cross-Validation

- Cross-validation is done using StratifiedKFold
- K is set to be 5

Outline

- Setup
  - Cross-validation
  - Class Balancing
  - Baselines
- Modeling
- Evaluation

Setup

❖Class Balancing

```
Home Loss: 0.2942
Draw: 0.2520
Home Win: 0.4537
```

- Class Balancing is done by loss weights
- First, The loss weight of N/C was tried
- Finally, we decided to not do any balancing

Outline

- Setup
  - Cross-validation
  - Class Balancing
  - Baselines
- Modeling
- Evaluation

Setup

❖Baselines

- Always Home Win: 45.5% Accuracy

- Bookmakers' Odds: Inverted and normalized

$$P(H) = \frac{\frac{1}{BO_H}}{\frac{1}{BO_H} + \frac{1}{BO_A} + \frac{1}{BO_D}}$$

Outline

- Setup
- Modeling
  - Chosen Algorithms
  - Comparison with baselines
  - Hyper-parameters tuning
- Evaluation

Modeling

❖Chosen Algorithms

- BladeChest

$$\mathbf{h}_{\text{blade}}\left(\mathbf{x}_h\right) = f\left(B\mathbf{x}_h\right)$$
$$\mathbf{a}_{\text{blade}}\left(\mathbf{x}_a\right) = f\left(B\mathbf{x}_a\right)$$
$$\mathbf{h}_{\text{chest}}\left(\mathbf{x}_h\right) = f\left(C\mathbf{x}_h\right)$$
$$\mathbf{a}_{\text{chest}}\left(\mathbf{x}_a\right) = f\left(C\mathbf{x}_a\right)$$

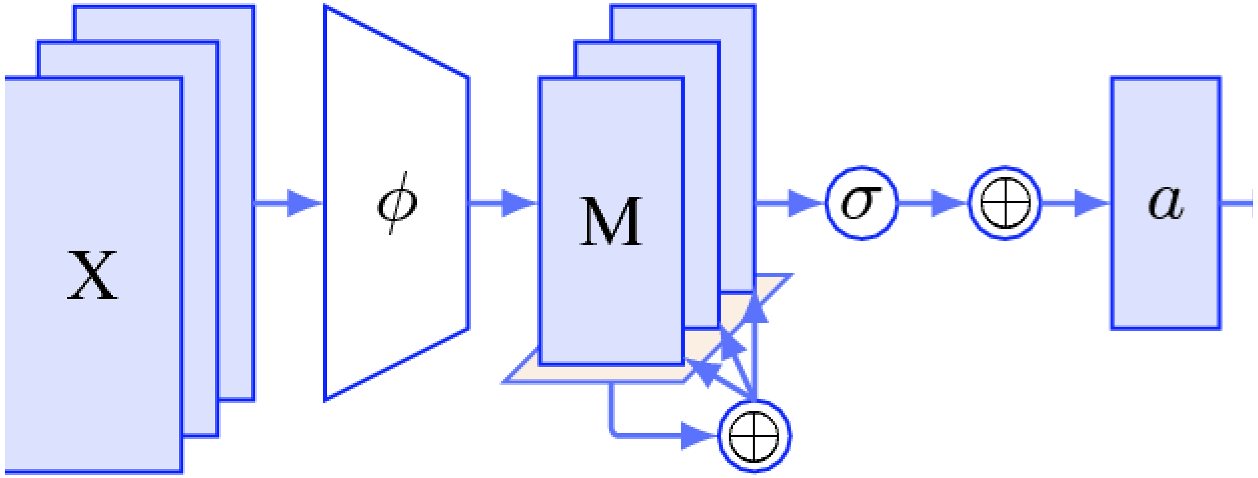
$$m_{h,a} = \mathbf{h}_{\text{blade}} \cdot \mathbf{a}_{\text{chest}} - \mathbf{a}_{\text{blade}} \cdot \mathbf{h}_{\text{chest}}$$

Modeling

❖ Chosen Algorithms

- Deep-Set Aggregation

$$\mathrm{MLP}_\theta\left(\sum_{a\in\mathcal{A}}\mathrm{MLP}_\phi\left(\mathbf{x}_a\right)\right)$$



Outline

- Setup
- Modeling
  - Chosen Algorithms
  - Comparison with baselines
  - Hyper-parameters tuning
- Evaluation

Modeling

❖ Comparison with Baselines

Always Home	45.37
Bookmakers	48.60
BladeChest	49.30
Set Aggregation	50.45

Outline

- Setup
- Modeling
  - Chosen Algorithms
  - Comparison with baselines
  - Hyper-parameters tuning
- Evaluation

Modeling

❖ Hyper-Parameter Tuning

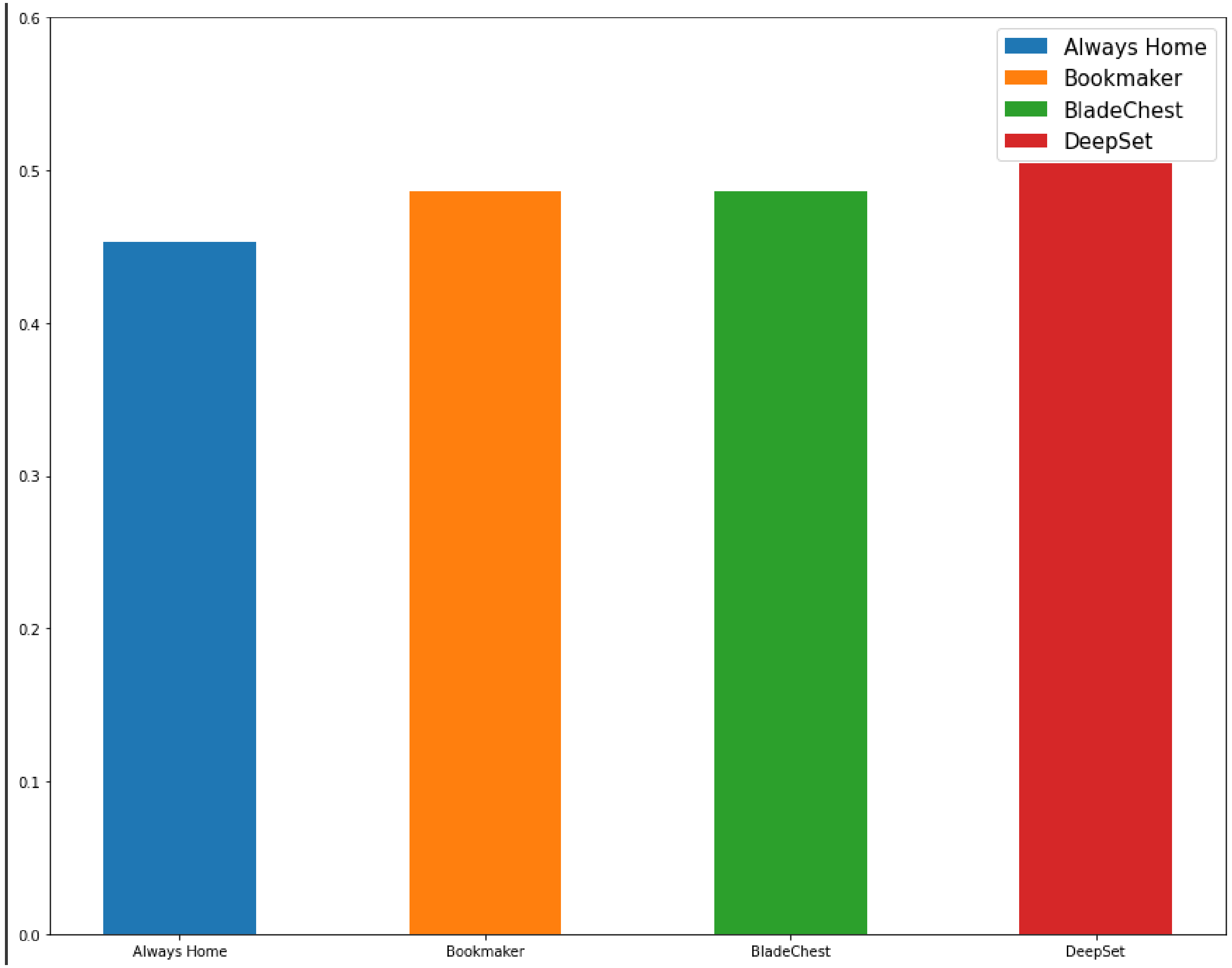
- ¬⅄Dropout: [0, **<u>0.25</u>**, 0.5]
- Dense sizes: [[6], **<u>[6, 6]</u>**]
- Epochs: [15, **<u>20</u>**]
- Learning Rate: [1e-3, **<u>3e-3</u>**]
- Player hidden sizes: [**<u>[8, 8]</u>**, [10, 14]]
- Team hidden sizes: [**<u>[8, 10]</u>**, [14, 14]]

Outline

- Setup
- Modeling
- Evaluation
    - Evaluation Metrics
    - Error Costs
    - Retuning to previous Phases
    - Discussion and future work

Evaluation

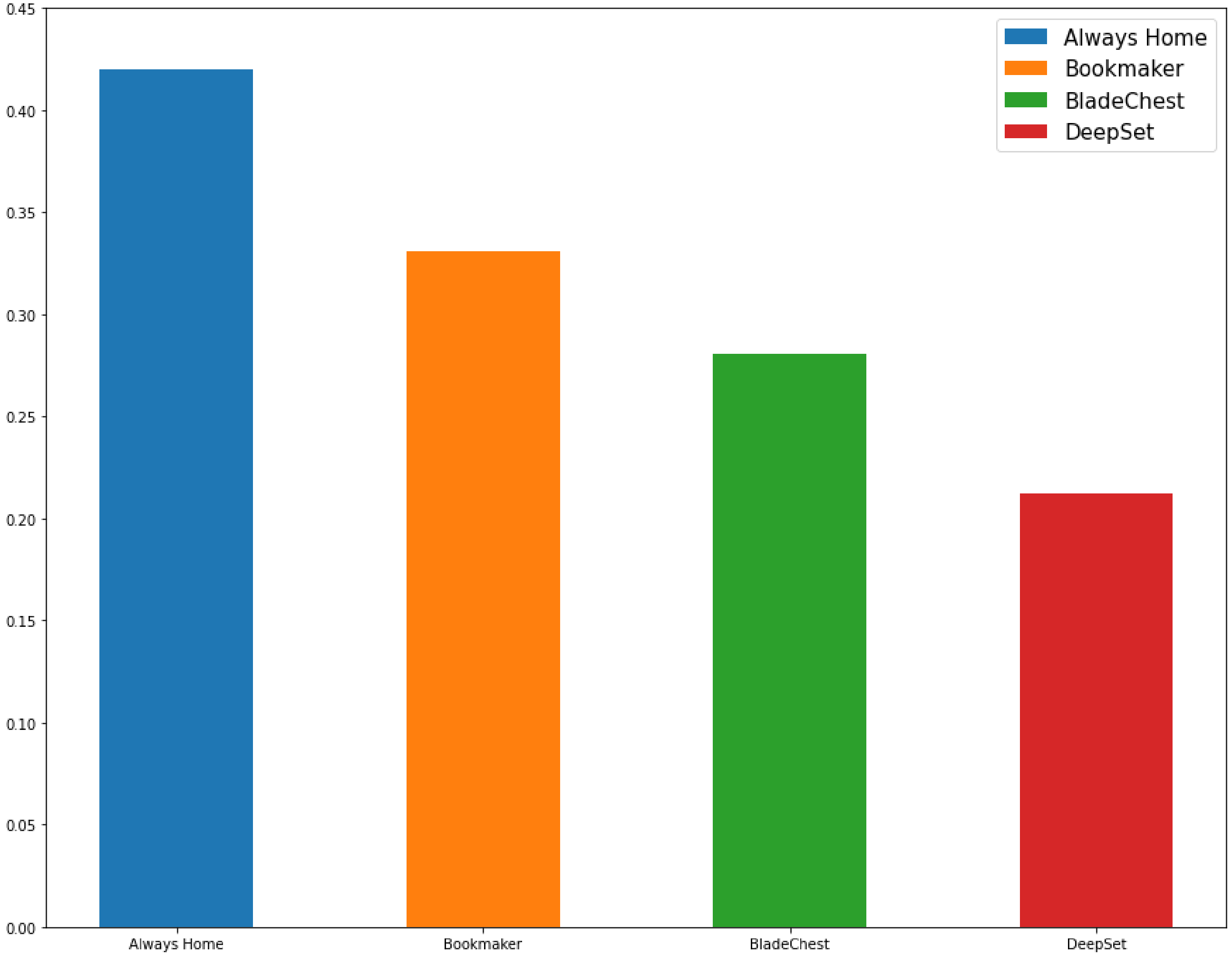
❖ Evaluation Metrics: Accuracy



Evaluation

❖ Evaluation Metrics: Ranked Probability Score

$$RPS = \frac{1}{r-1}\sum_{i=1}^{r-1}\left(\sum_{j=1}^{i}(p_j - y_j)\right)^2$$



Outline

- Setup
- Modeling
- Evaluation
  - Evaluation Metrics
  - Error Costs
  - Retuning to previous Phases
  - Discussion and future work

Evaluation

❖ Error Costs

- The nature of football match outcome prediction is non-critical
- Costs would be in betting or revealing bookmakers intentions
- Always betting the minimum odd results in loosing money!

Outline

- Setup
- Modeling
- Evaluation
  - Evaluation Metrics
  - Error Costs
  - Retuning to previous Phases
  - Discussion and future work

Evaluation

❖ Returning to Previous Phases

- Built-in Normalization
- Omitting the only categorical feature
- Not Binning numerical features
- Changing rolling average window to 4

Outline

- Setup
- Modeling
- Evaluation
  - Evaluation Metrics
  - Error Costs
  - Retuning to previous Phases
  - Discussion and future work

Evaluation

❖ Discussion and Future Work

The Blade Chest model tries to capture the **interactions between teams** through two feature vectors each encompassing the **offensive and the defensive** strength of teams which is **similar to the real world** analysis. The **biggest drawback for this model is its ignorance towards players and team lineup**.

The Set Agg model's strength lies in its ability to **include team players** and their previous performances in both spatial and temporal contexts. However, the **temporal aspect of modeling could also be modeled through machine learning approaches**.

One **suggestion to improve** the Set Agg's ability to **capture the evolution** of players through time is to use a **temporal graph structure** wherin each player has a chained structure spanning through time. The spatial context could easily be modeled with graph structure, forming a spatio-temporal message passing graph.

THANK YOU!