Deep learning plays Handball

•••

Pere Urbon-Bayes
pere.urbon@gmail.com

Agenda

- How can AI help Handball.
- What others are doing already.
- Let’s do play classification all together.

What others are already doing?

or other invasion sports.

How can AI help Handball teams?

Handball teams, even professional ones, don’t have the similar budgets as other popular sports such as Football or Basketball.

- Reduced budgets limit the headcount for an sport analyst team.
- Human analysis in invasion sports is a known biased activity that require big amounts of time to do successfully.
- Current technology advances have permitted teams in professional sports to gather detailed information about matches, as well as trainings. See HBL.

Example in Germany? Handball Bundesliga

- Since the season 2019/20, the HBL uses KINEXON technology to do indoor track for players and balls during games.
- Data is gathered using an indoor positioning system.
- Derived an HPI index, to track player performances.
- Example attributes collected
  - Player location in the field.
  - Ball location in the field.
  - As derived attributes: ball throw speed
  - Goal technology: Throw landing area in the goal

How can AI help Handball teams?

With a powerful ML/AI system, the handball teams could:

- Get a better understanding of their opponent plays/strength, without relying on expensive, and not accurate human analysis.
- Estimate which offensive plays have more success (*expected possession value*).
- Better video segmentation, facilitating search.
- More accurate play segmentation, facilitating search (as well!)
- .....

Why bother with AI in Handball?

or other invasion sports.

Play classification

In this category, the aim would be to label the play in analysis.

Having multiple degrees of complexity, this family of problems could range from a multi-classification problem (most difficult) to the most basic binary classification problem (organized or not).

Play classification

Mures et al. developed PlayNet, a system used in the spanish handball leagues to help autocontrol the video cameras.

Such a system would run in low resources and need quick inference times, so it uses Kalman embeddings and traditional ML (RandomForest / KNN).

Classifications are focused to help camera movements, left_attack, right_penalty, etc..

Play classification

Plays in invasion games, specially in Handball or Basketball, follow a given choreography, usually involving semantically differentiated movements:

- Left or Right wings transitioning to secondary pivots.
- Pivot positioned in one of the sides, and used by the first line transitions to create a superiority.

In their research Matej Perse et al. develop a system that using video recording and gaussian functions on player trajectories, velocities, distance and accelerations was able to create a semantic encoding for basketball plays.

While the system involve multiple expensive steps, including manual annotations, the core principle is very powerful in the area.

Expected possession value

Calculating a possession value or giving a possession result is the process of estimating the success of a game possession.

This task could be achieved either by determining a team strength score, scoring probabilities or the simplified version of determining the result via a binary classification problem.

Expected possession value

Adams et al. develop a model to predict an expected goal score, with the idea to help understanding scoring opportunities in handball.

Based on positional data, and derived attributes related to the action (throwing angle, distance to the goal, ...) the authors compare NN and more traditional ML for the job.

LGBM and CatBoost, both gradient boosting methods, performed the best on weighted f1-score.

Note, the was develop with/for the TBV Lemgo Lippe team in the HBL.

Expected possession value

Müller et al. created PIVOT, a framework to predict the probability of an scoring event in the last T seconds.

The authors use only positional data, and derived attributes.

In PIVOT, the creators benchmark multiple NN architectures, Time series transformers (TST), 3 layer LSTM, 3 layer FCN. Being the best on the experiments the TST with AUC 0.9.

Note, the was develop with/for the SG Flensburg-Handewitt team in the HBL.

Challenges in Sports AI

Challenges in Sports AI

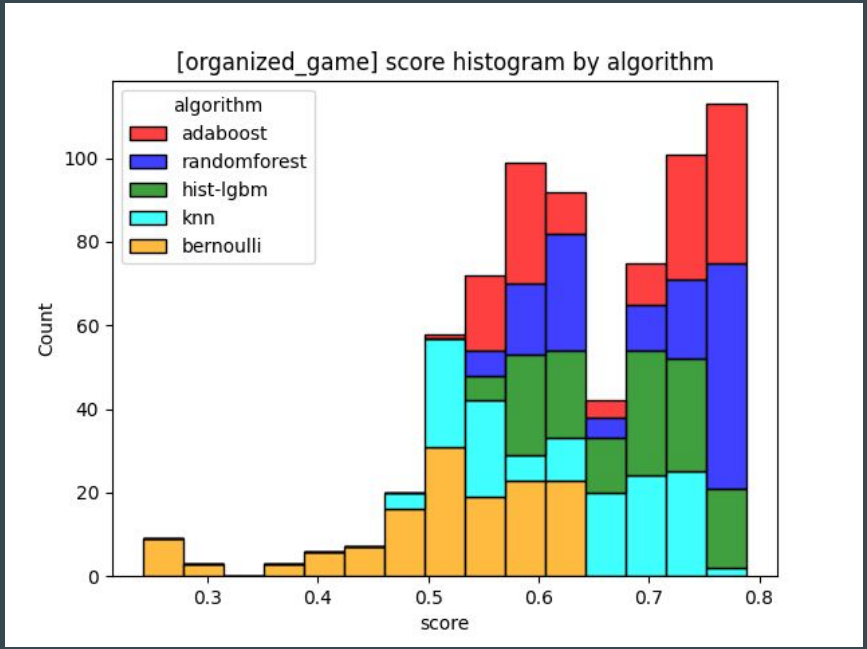
- Datasets are highly unbalanced (by analysis property)
  - There is the need to do re-sampling.
  - Studies indicate conflicting reports on the benefits of this techniques.
  - Generative AI is an unexplored field.
- Access to high quality datasets is limited to top ligas (male HBL mostly).
- Available datasets are not diverse enough
  - Male vs female
  - Defensive and offensive strategies vary by playing school (east europe, germany, spain, others)

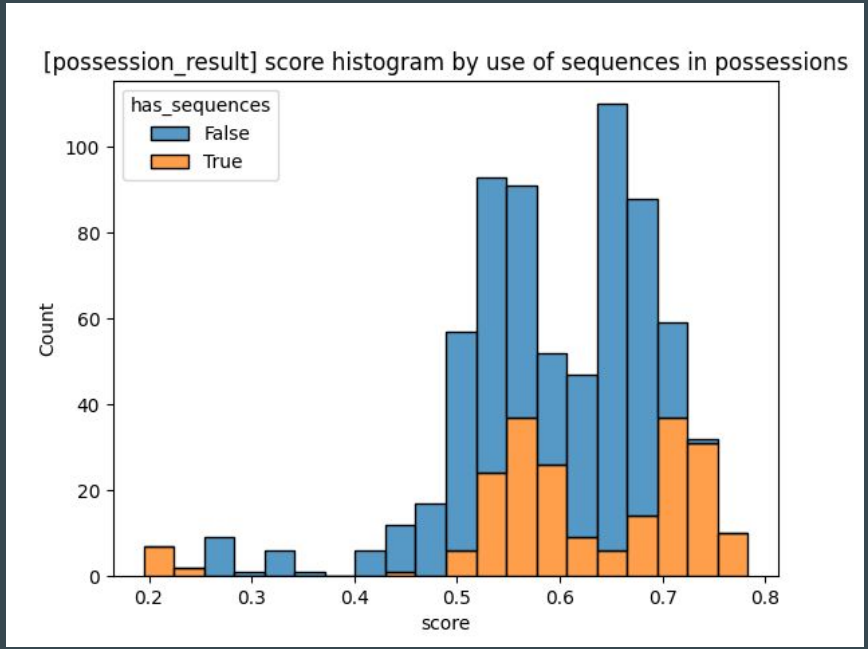
Reporting results

- Because of the datasets characteristics, generally studies report results using **f1-score**.
- However, there are many studies where results are only reported using AUC, without adding more context, causing confusion on the final performance.

Let’s do play-classification together!

Some (other) experimental results





References

The resources presented here can be found:

- Demo code with models https://github.com/purbon/bbuzz-handball/tree/main/playmaker
- Presentation: https://github.com/purbon/bbuzz-handball

The dataset used for the training can’t be shared yet, however will be published later this year.

Thank you very much! Questions?