Women's football analyzed: interpretable expected goals models for women

Lotte Bransen[1,2] , **Jesse Davis**[1]
[1]KU Leuven, Department of Computer Science & Leuven.AI, Belgium
[2]SciSports, The Netherlands
lotte.bransen@kuleuven.be, jesse.davis@kuleuven.be

Abstract

Technical data such as event or optical tracking data from men's football (soccer) matches have been extensively analysed using techniques from AI on a variety of different levels. However, there has been very little analysis of the women's game. In this work we take an initial step towards analysing professional women's football. Using event data covering a number of seasons from the top women's leagues, we perform two analyses. First, we perform an exploratory analysis by computing several technical indicators (e.g., goal scoring rates over the season, conversion rates, shot locations) and then compare and contrast them to the indicators for comparable men's leagues and find several intriguing differences. Second, we assess whether xG models on one gender are applicable to data from a different gender.

1 Introduction

The women's football game has made huge advances in recent years. When the first official FIFA Women's World Cup took place in 1991, it featured matches that lasted only 80 minutes and the final was not even shown on TV [Harris, 2015]. Just 28 years later, more than 1 billion viewers watched the 2019 World Cup final between the Netherlands and the USA [FIFA, 2019]. Concurrently, financial investment in the women's club realm has increased the number of players who are able to play professionally. Recently, UEFA announced a new Women's Champions League format with more money and more teams [UEFA, 2021]. The net effect is an increased level of competitiveness.

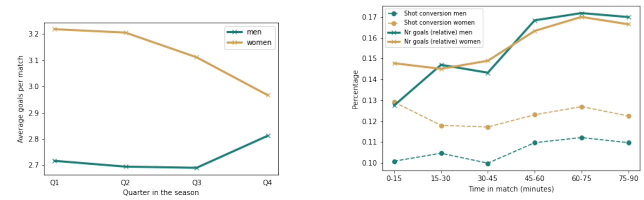
Despite its increased popularity, there has been less analysis of data related to the women's game. Most existing studies focusing on physical aspects of the game. [Pedersen *et al.*, 2019] state that the differences in the style of play between women and men are mostly due to the physical differences like endurance, kicking velocity, height, speed and foot length. [Bradley *et al.*, 2014] discovered that male players covered more distance at a higher speed and that these differences were bigger in the second half. [Cardoso de Araújo *et al.*, 2020] found the largest differences in sprints, jumps and intermittent endurance between female and male football players. In contrast, analyzing the technical aspects of the game has received less attention. [Worville, 2020] compared the shots taken in the English Women's Super League and the Premier League and [Sakellaris, 2017] compared the number of goals in national team matches and concluded that in most cups women score more goals per match. [Pappalardo *et al.*, 2021] performed a more extensive comparison of women's and men's football in World Cup matches. They concluded that men shoot from further away and have higher pass accuracies, while women regain possession quicker and prefer short passes over long balls. Furthermore, they use several indicators to train a model that predicts whether these represent a women's or men's match. Similarly, [Casal *et al.*, 2021] saw more accurate passes in Spanish men matches than in women matches. [Garnica-Caparrós and Memmert, 2021] analyzed event data from the men's 2016 and women's 2017 European Championships. They trained models to predict a player's gender based on features such as the number of times a player (un)successfully performed a certain type of action.

This research takes a first in-depth analytical look using machine learning at the technical data that is now being collected from professional women's matches. Specifically, we focus on shots as the fundamental objective of football is to score more goals than your opponent, and as Johan Cruijff famously said: "you can't score if you don't shoot." A natural way to analyse shots and shot behaviour is through the lens of the well-known expected goals (xG) metric, which gives the probability that a shot will yield a goal. This metric helps cope with the fact that evaluating players and teams based on the number of goals scored can be undesirable because goals are relatively rare and subject to random fluctuations and luck (e.g., deflections). In contrast, xG quantifies the quality of the chances created which enables better understanding the performance of a team or player. Expected goals has become a relatively mainstream metric and is discussed on TV shows like BBC's Match of the Day and has even been used in the popular Football Manager computer game [SciSports, 2020].

Needless to say, expected goals has been extensively researched [Pollard and Reep, 1997; Lucey *et al.*, 2014; Caley, 2015; Rathke, 2017; Robberechts and Davis, 2020; Madrero Pardo, 2020; Anzer and Bauer, 2021]. However, as far as we know, no research has focused on the women's game. This could be because until recently, little data about the women's game was collected and even less of it is pub-

League	M/F	17/18	18/19	19/20	20/21
WSL	F	606	2292	2005	2488
Division 1	F	2147	2737	2048	2224
Bundesliga	F	2859	3298	3356	2301
Primera Div.	F	2041	2382	2917	4685
NWSL	F	2865	2621	2698	25
Premier League	M	8401	8505	8287	6855
Ligue 1	M	8274	8180	6008	6874
Bundesliga	M	6884	7288	7111	5981
Serie A	M	8849	9375	9632	7166
Primera Div.	M	8090	8122	7485	6001

Table 1: Number of shots in our data set per league per season. Some of the 19/20 seasons are incomplete due to the Covid-19 pandemic. The 2020 NWSL season only contains 25 shots because it was cancelled due to the pandemic.



(a) The average number of goals per match over the season. The number of goals per women's match (orange) declines over the season whereas men's matches (blue) sees an increase near the end of the season.

(b) The average number of goals and the shot conversion rate for each 15 minute interval in a match. Women (orange) tend to score more and have a higher shot conversion rate than men (blue) at the start of the match.

Figure 1: Distribution of goals for women's (orange) and men's games per (a) each quarter of the season, and, (b) each 15 minute interval of a match.

licly available, with a notable exception being StatsBomb releasing some data about the English WSL. While the data contains xG values for taken shots, it is too small to enable training an accurate xG model or performing a more in-depth analysis. In this paper, we fill this gap by analysing multiple seasons of event data from five top professional women's leagues. Moreover, we will contrast the shooting behaviour between women's and men's competitions. Specifically, we will address the following questions:

1. What, if any, are the differences on a match and competition level between women's and men's leagues?
2. What, if any, are the differences in shot selection and conversions between top women's and men's leagues?
3. Are xG models learned on data from women's matches transferable to men's matches and vice versa?

Our analysis produced a number of intriguing findings such as that women's professional matches have less extra time than men's matches, women convert a higher percentage of their shots than men, and women both head closer to the goal and score more often from headers. Our analysis of xG models finds that they are indeed transferable across matches of different genders though some specific shot types are valued differently. These observations greatly enhance our understanding of the way that the women's game is being played.

2 Data

Our analysis considers the five strongest women's and men's football leagues. The women's leagues included are: the English WSL, American NWSL, German Frauen Bundesliga, French Feminine Division 1 and Spanish Primera Division Femenina. The considered men's leagues are the English Premier League, Spanish Primera Division, French Ligue 1, German Bundesliga and Italian Serie A. The data set contains event stream data encoded in the SPADL format [Decroos *et al.*, 2019] from matches in the 2017 and 2017/2018 seasons until the 25th of April 2021 and was provided by SciSports.[1] In total, we have information about 9,076 matches of which 2,100 are women's matches and 6,976 are men's matches.

We focus on open play shots and hence omit penalties, direct freekicks and own goals. Table 1 shows the number of shots for each league in each considered season. As women's leagues typically have fewer teams, these leagues have fewer matches and shots per season. Furthermore the 2019/20 and 2020 seasons contain less data as the Covid-19 pandemic led to some matches being cancelled in these seasons. Due to the unfinished 2019/20 Primera Division Femenina, two extra teams were added to the league for the ongoing 2020/21 season. Finally, data for some matches in this league is also missing for the 2017/18 and 2018/19 seasons resulting in a lower number of shots for these seasons.

3 Observations

We discuss several interesting insights obtained from our data set in terms of differences between women's and men's football on two levels; (1) goals in matches and competitions, and (2) shot behaviour and outcomes during matches.

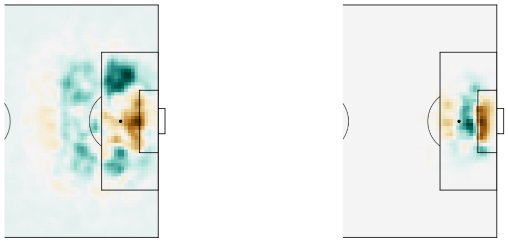
3.1 Matches and competitions

First, we explore differences between the two games regarding shots across matches and competitions.

Women score more goals at the start of the season We analyze matches from complete seasons (i.e., excluding the Covid-19 seasons (2019/20 and 2020) and the ongoing 2020/21 seasons) to see how the number of goals varies over the course of a season. We divide each season into quarters based on the number of matches in the season and compute the number of goals scored per quarter. Figure 1(a) shows that more goals are scored in men's matches near the end of the season. However, this is different in women's leagues, where more goals are scored at the start of the season.

Women's leagues contain more less-competitive matches In general, larger goal differences arise in women's matches, with 18% of the matches having a goal difference of more than 3 goals opposed to 6% in the men's matches. This likely arises because the women's game is still rapidly evolving and

[1] https://www.scisports.com



(a) Shots with the foot. Women shoot more often directly in front of goal inside the box whereas men shoot more from "Robben" locations and just outside the box.

(b) Headed shots. We observe that women tend to head closer to the goal than men.

Figure 2: Difference in proportion of shots with the (a) foot and (b) head from each location. Orange (blue) areas indicate more shots in women's (men's) matches.

an increasing number of players can nowadays make a living from playing football.

Draws are rare in the women's game In our data set, 17% of the women's matches end in draws versus 25% of the men's matches. However, women's matches have more away wins and a slightly higher number of home wins.

Women matches have less extra time We observe that women's matches tend to have on average around 37 seconds less time added to the regular 90 minutes of play. This includes time added in both the first and second half. This is quite a big difference, given that women's matches have about 5 minutes extra time. The two English leagues have the longest extra time with the WSL having 6.8 and the Premier League 6.5 added minutes.

3.2 Shot behaviour

Next, we explore differences in shot behaviour in terms of conversion rates and where shots arise.

Women have a higher shot conversion rate On average 12.2% of the women's shots find the net opposed to 10.7% in men's matches. The exception is the American NWSL where the conversion rate is 9.9%.

Shot conversion rates change throughout the course of the match Research showed that men tend to score more goals near the end of a match [Armatas *et al.*, 2007]. We investigate whether this claim also holds in the women's game. Figure 1(b) shows the average number of goals and the shot conversion rates for both games for every fifteen minutes in a match. The number of goals in both games shows a similar pattern throughout the match, with more goals near the end of the match. The shot conversion rates seem to be stable throughout the match.

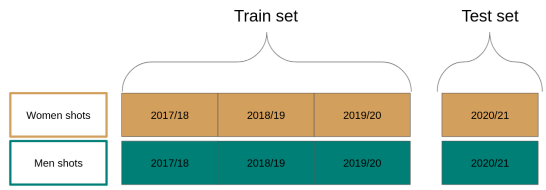


Figure 3: We split our split our data sets into a train and a test set using a temporal split and train our xG models on three different train sets; a women only, a men only and a combined set.

Fewer "Robben" shots in the women's game Figure 2(a) shows the difference in proportion of shots with the foot from each location with orange (blue) areas indicating more shots in women's (men's) matches. Women tend to shoot from locations with a smaller angle to the goal and in general from closer distances to the goal. Interestingly, men shoot more often than women from the so-called "Robben" location[2], inside the box near the corner of the box where an inverted winger can shoot and curl the ball into the far corner of the goal.

Women head closer to the goal and score more often from headers Figure 2(b) shows the difference in proportion of headed shots from each location with orange (blue) areas indicating more shots in women's (men's) matches. Generally, women head closer to the goal than men do with the main difference arising inside the 5 meter box. Women score more often from headers with 15.7% of their headers yielding a goal opposed to 12.7% in the men's game.

4 xG Models

We address the question: are xG models learned on data from women's matches transferable to men's matches and vice versa? To answer this question we consider three datasets:

Women only contains only shots from women's matches.

Men only contains only shots from men's matches.

Combined contains all shots.

We split each data set into a train and test set using a temporal split: the train set has matches from the 2017/18, 2018/19 and 2019/20 seasons and the test set has matches from the 2020/21 season. Figure 3 visualizes the breakdown of our data sets. This split results in a total number of 157,346 shots for training (36,868 women, 120,478 men) and 44,589 shots in the test sets (11,719 women, 32,870 men).

4.1 Challenges

Comparing models trained on different data sets is not straightforward for several reasons:

Rapid evolution women's game Due to the rapid evolution of the women's game, historical shot data might not be representative for the game as it is played today. Therefore, we do not go further back in time than the 17/18 season and the 2017 season for the NWSL.

[1]https://github.com/TomDecroos/matplotsoccer

[2]Arjen Robben was one of the first prominent inverted wingers and he frequently shot from this area of the pitch.

Covid-19 Due to the Covid-19 pandemic, our data set contains many incomplete seasons of data. Our analysis indicates that the number of goals scored varies over the course of a season. Therefore we need to correct for the fact that our data set contains fewer shots from the last quarter of the seasons due to the pandemic.

Competition set up Women leagues have fewer teams, and thus matches and shots. We do not want the combined model to overfit on the men's shots. Therefore need to correct for the difference in the number of women and men shots such that it contains the same number of shots from both games while still preserving the ratio of shots by the home team to the away team.

4.2 Data set sampling

To overcome the aforementioned challenges we use sample weighting to ensure that some shots in our data set will be more important during training than others. We assign weights to matches, not to individual shots, to ensure that all shots within a match are weighted similarly. We weight samples as follows:

1. **Women shots** Given that we have more shots by men than women, we weight the shots by women in our train set over three times more than our shots by men. Thus we have the same number of women and men matches in our train set.
2. **Matches per quarter** We weight the matches in such a way that we have an equal number of matches from each quarter of the season in our data set.
3. **Home and away matches** Once we have corrected for the matches per quarter, some teams may have a big difference in the number of home and away matches. We correct for this by weighting shots per team per home/away match based on its home/away ratio in our train set.

4.3 xG models

On each data set, we train two xG models:

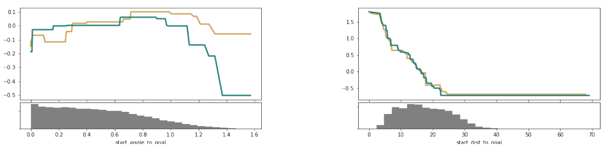
Generalized Additive Models (GAMs) are interpretable models that have been successfully applied to football data [Decroos and Davis, 2020]. We use the `interpret-ml`[3] package to train the GAMs using their Explainable Boosting Classifiers [Nori *et al.*, 2019] that use a boosting procedure to learn the feature functions of the GAM and to select pairwise interaction terms.

XGBoost [Chen and Guestrin, 2016] which learns an ensemble of trees, which have yielded excellent performance on xG and other football tasks [Anzer and Bauer, 2021; Madrero Pardo, 2020; Decroos *et al.*, 2019]. We use `xgboost`'s[4] XGBClassifier tor train the model.

We refer to the six models based on the considered data and type of model: TotalGAM, WomenGAM, MenGAM, TotalXGB, WomenXGB and MenXGB.

[3]https://github.com/interpretml/interpret

[4]https://xgboost.readthedocs.io/



(a) Angle between the shot location and center of the goal

(b) Distance to the goal

Figure 4: Partial dependence plots for the angle to the goal and the distance to the goal for the WomenGAM (orange lines) and the MenGAM (blue lines). The histograms on the bottom show the features' distributions of values in the combined test set.

We describe each shot using 12 features such as the shot's location, the type of assist (e.g. from dribble, from cross, possession change), the body part used (e.g. foot, head), the scoreline, and time in the match as calculated by the `soccer-xg`[5] package. We predict whether the shot yields a goal and are mostly interested in the probabilities that the models produce which define the expected goals values. We optimize the models' hyperparameters using a gridsearch with 5-fold cross-validation.

5 Results

This section describes the results of our experiments. Firstly, we inspect our GAMs to gain insight into whether a feature's importance varies between the women's and men's games. Secondly, we analyze the performance of our six models on a per data set basis to help us answer the question whether we can transfer models trained on one gender to the other. Thirdly, we analyze the predictions made by the models on different types of shots and show some examples where the models clearly do not agree. Finally, we analyze the average xG value of a shot and the percentage of shots with a low xG value per league in the 2020/21 season.

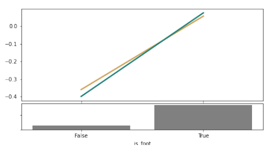
5.1 Model inspection

The additive nature of GAMs permit analyzing the impact of each feature on the predictions by using partial dependency plots. We compare the partial dependence plots for our WomenGAM model to our MenGAM model to analyze the similarities and differences.

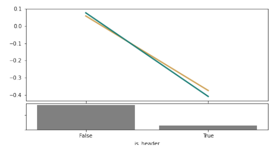
Shot location Figure 4 shows the partial dependence plots for two important features regarding the shot's location: the angle and distance to the goal. Interestingly, the angle to goal has a differing impact in both models with larger angles negatively impacting the probability of a goal in the men's model while having a negible impact in the women's model. In contrast, the relationship between the distance to the goal is similar in both models.

Assist type and body part Figure 5 shows the partial dependence plots for a number of features describing the shot's assist and the body part used to execute the shot. We can summarize the insights from these plots as follows:

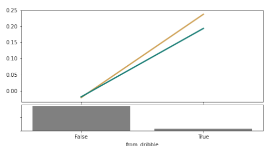
[5]https://github.com/ML-KULeuven/soccer_xg



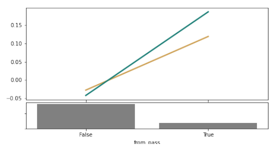
(a) Foot shot



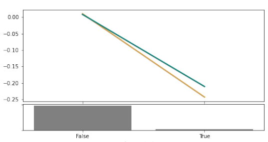
(b) Header



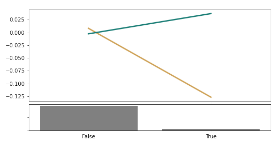
(c) Shot following a dribble



(d) Shot directly from a teammate's pass



(e) Shot directly from a teammate's set piece

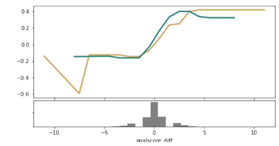


(f) Shot directly from a teammate's cross

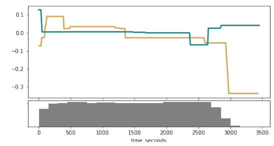
Figure 5: Partial dependence plots for features describing the body part used for the shot and action preceding the shot for the WomenGAM (orange lines) and the MenGAM (blue lines). The histograms on the bottom show the features' distributions of values in the combined test set.

1. Shooting with the foot has a slightly smaller positive impact in the WomenGAM whereas a header has a slightly less negative impact in the WomenGAM.
2. Shots following a dribble get a higher xG value in the WomenGAM, while shots following a teammate's pass get a higher xG value in the MenGAM.
3. Both models assign lower xG values to shots following a set piece, with the WomenGAM producing a larger negative impact.
4. The WomenGAM assigns lower xG values to shots directly from a cross, whereas this has a slight positive effect in the MenGAM.

Game state Figure 6 shows the partial dependence plots for two features describing the game state when the shot was taken: goal difference and time in the current half. In both models, higher positive goal differences result in higher xG values. This finding agrees with past work [Caley, 2015], which was explained by the fact less defensive pressure will be applied by the opponent when leading. This feature could also be a proxy for team strength as better teams tend to score more. Slightly larger goal differences have a bigger impact in the MenGAM than the WomenGAM. The WomenGAM's xG values tend to decrease over time, apart from the first minutes when fewer shots are taken in general. However, a different pattern arises in the MenGAM where xG values drop near the end of the half before increasing again in extra time.



(a) Goal difference



(b) Number of seconds in the current half

Figure 6: Partial dependence plots for the features describing the game state for the WomenGAM (orange lines) and the MenGAM (blue lines). The histograms on the bottom show the distribution of values for the feature in the combined test set.

Model/Test set	Women only	Men only	Combined
WomenXGB	0.09705	0.08677	0.08947
WomenGAM	0.09714	0.08680	0.08952
MenXGB	0.09675	0.08636	0.08909
MenGAM	0.09722	0.08674	0.08950
TotalXGB	0.09675	0.08635	0.08908
TotalGAM	0.09690	0.08654	0.08926

Table 2: Brier scores for all six models on each of the three test sets.

5.2 Model evaluation

We estimate the xG values for the shots in our three test sets using the three GAM models and three XGBoost models. Table 2 shows the Brier Score each model achieves on each data set. We use this metric as the most important criteria is having a well-calibrated model. All models achieve similar performance. Figure 7 shows the calibration plots for the six models on the combined test set. All models are well-calibrated for the more commonly occurring low xG shots and, predictably, struggle for the rarer big chances that yield higher xG values.

5.3 Individual shots with big xG differences

We investigate the predictions made by the different models on individual shots with a particular emphasis on identify-

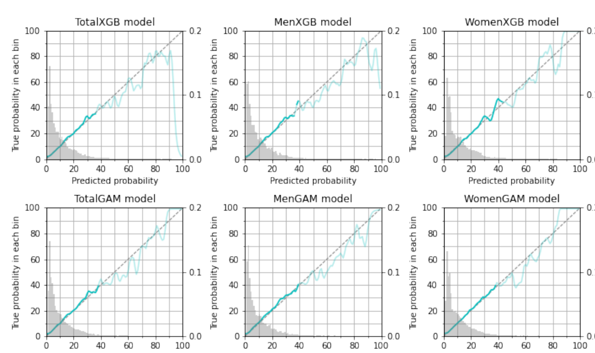


Figure 7: The calibration plots for all six models on the combined test set. In general, all models are well-calibrated for the probabilities on the most commonly occurring shots.

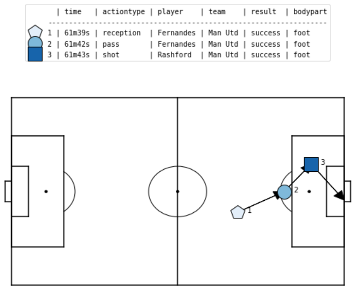


Figure 8: Marcus Rashford's 62nd minute goal from the "Robben" location in the 2-1 win against Brighton on April 4th, 2021.

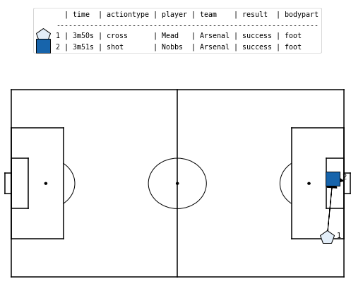


Figure 9: Tip-in by Jordan Nobbs following a Beth Mead cross in the 4-0 Arsenal win against Everton on December 20th, 2020.

ing shot types where the models produce differing xG values. The women-based models and men-based models disagree most for the so-called "Robben" shots, tip-ins from short distance following a cross, and headers following a set piece.

"Robben" shots The models trained on men's shots assign higher values to "Robben" shots than the models trained on women's shots. Figure 8 shows an example of such a shot by Marcus Rashford. This shot received xG values of 0.069 (MenGAM) and 0.087 (MenXGB). However, the models trained on women's data produced lower values: 0.037 (WomenGAM) and 0.066 (WomenXGB).

Tip-ins following a cross The models trained on women's data assign higher values to tip-ins close to the goal than the models trained on men's data do. Figure 9 shows an example of such a goal by Jordan Nobbs which received xG-values of around 0.60 from the women's models (0.586 from WomenGAM and 0.608 from WomenXGB). However, the men's models assigned a much lower xG value of around 0.48 (0.478 from MenGAM and 0.476 from MenXGB).

Headers following a set piece The models trained on men's data assign higher xG values to headers following a set piece than the women's models do. Figure 10 shows an example of such a header by Brighton & Hove Albion's Victoria Williams. This opportunity was valued by the women's models at 0.196 (WomenGAM) and 0.359 (WomenXGB), but received substantially higher xG values of 0.654 (MenGAM) and 0.614 (MenXGB) from the men's models.

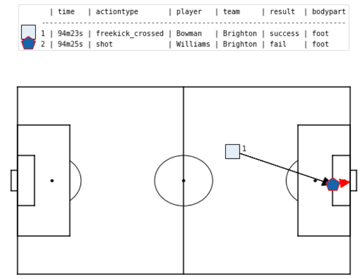


Figure 10: Late Victoria Williams miss in the 95th minute following a freekick in the 3-1 loss against Tottenham on December 6th, 2020.

League	M/F	avg xG	% low xG (<0.02)
WSL	F	0.128	3.5%
Division 1	F	0.134	3.6%
Bundesliga	F	0.127	3.1%
Primera Div.	F	0.126	4.3%
Premier League	M	0.111	10.3%
Ligue 1	M	0.105	14.0%
Bundesliga	M	0.110	11.2%
Serie A	M	0.112	10.3%
Primera Div.	M	0.110	13.3%

Table 3: The average xG value per shot and the percentage of low xG shots per league. We observe lower average xG values and a higher percentage of low xG shots in the men leagues.

5.4 xG values across leagues and gender

Table 3 shows the average xG value per shot (based on the XGBoost model per gender) in our test set for each league, where we exclude the NWSL due to its small sample size. Women tend to take shots with higher chances of yielding a goal and men take more speculative shots, that is, those having a low chance of yielding a goal.

6 Conclusion

This paper performed an extensive analysis of women's football shots. We identified interesting observations such as the fact that women tend to shoot from different locations than men, have a higher shot conversion rate and their goals are differently distributed across the season. We trained six different xG models on different data sets with different machine learning algorithms and found that, in general, models from one gender are applicable to shots from the other. However, when inspecting the models and shots, some interesting differences arose in terms of what features are important and how the models value certain types of shots.

Acknowledgements JD is partially supported by KU Leuven Research Fund (C14/17/070) and the Flemish Government under the "Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen" programme.

References

[Anzer and Bauer, 2021] Gabriel Anzer and Pascal Bauer. A goal scoring probability model for shots based on synchronized positional and event data in football (soccer). *Frontiers in Sports and Active Living*, 3:53, 2021.

[Armatas *et al.*, 2007] Vasilis Armatas, A. Yiannakos, and P. Sileloglou. Relationship between time and goal scoring in soccer games: Analysis of three world cups. *International Journal of Performance Analysis in Sport*, 7:48–58, 05 2007.

[Bradley *et al.*, 2014] Paul S. Bradley, Alexandre Dellal, Magni Mohr, Julen Castellano, and Anna Wilkie. Gender differences in match performance characteristics of soccer players competing in the uefa champions league. *Human Movement Science*, 33:159–171, 2014.

[Caley, 2015] Michael Caley. Premier League Projections and New Expected Goals. Cartilage Free Captain, 2015.

[Cardoso de Araújo *et al.*, 2020] Maithe Cardoso de Araújo, Christian Baumgart, Christian T. Jansen, Jürgen Freiwald, and Matthias W. Hoppe. Sex differences in physical capacities of german bundesliga soccer players. *Journal of Strength and Conditioning Research*, 34:2329–2337, 2020.

[Casal *et al.*, 2021] Claudio Casal, José Losada, Rubén Maneiro, and Antonio Ardá. Gender differences in technical-tactical behaviour of la liga spanish football teams. *Journal of Human Sport and Exercise*, 16(1):37–52, 2021.

[Chen and Guestrin, 2016] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. pages 785–794, 08 2016.

[Decroos and Davis, 2020] Tom Decroos and Jesse Davis. Interpretable Prediction of Goals in Soccer. In *AAAI 2020 Workshop on Artificial Intelligence in Team Sports*, 2020.

[Decroos *et al.*, 2019] Tom Decroos, Lotte Bransen, Jan Van Haaren, and Jesse Davis. Actions Speak Louder Than Goals: Valuing Player Actions in Soccer. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1851–1861. ACM, 2019.

[FIFA, 2019] FIFA. FIFA Women's World Cup 2019™ watched by more than 1 billion. FIFA, 2019.

[Garnica-Caparrós and Memmert, 2021] Marc Garnica-Caparrós and Daniel Memmert. Understanding gender differences in professional european football through machine learning interpretability and match actions data. *Scientific Reports*, 11, 2021.

[Harris, 2015] Nick Harris. Number-crunched: The rise and rise of the Women's World Cup. Sporting Intelligence, 2015.

[Lucey *et al.*, 2014] Patrick Lucey, Alina Bialkowski, Mathew Monfort, Peter Carr, and Iain Matthews. Quality vs. Quantity: Improved Shot Prediction in Soccer Using Strategic Features from Spatiotemporal Data. In *MIT Sloan Sports Analytics Conference*, 2014.

[Madrero Pardo, 2020] Pau Madrero Pardo. Creating a model for expected goals in football using qualitative player information. Master's thesis, Universitat Politècnica de Catalunya, 2020.

[Nori *et al.*, 2019] Harsha Nori, Samuel Jenkins, Paul Koch, and Rich Caruana. Interpretml: A unified framework for machine learning interpretability. *arXiv preprint arXiv:1909.09223*, 2019.

[Pappalardo *et al.*, 2021] Luca Pappalardo, Alessio Rossi, Giuseppe Pontillo, Michela Natilli, and Paolo Cintia. Explaining the difference between men's and women's football. *arXiv preprint arXiv:2101.01662*, 2021.

[Pedersen *et al.*, 2019] Arve Vorland Pedersen, Ingvild Merete Aksdal, and Ragna Stalsberg. Scaling demands of soccer according to anthropometric and physiological sex differences: A fairer comparison of men's and women's soccer. *Frontiers in Psychology*, 10:762, 2019.

[Pollard and Reep, 1997] Richard Pollard and Charles Reep. Measuring the effectiveness of playing strategies at soccer. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 46:541 – 550, 01 1997.

[Rathke, 2017] Alex Rathke. An examination of expected goals and shot efficiency in soccer. *Journal of Human Sport and Exercise*, 12, 01 2017.

[Robberechts and Davis, 2020] Pieter Robberechts and Jesse Davis. How Data Availability Affects the Ability to Learn Good xG Models. In *International Workshop on Machine Learning and Data Mining for Sports Analytics*, 2020.

[Sakellaris, 2017] Dimitrios Sakellaris. The In-Game Comparison Between Male and Female Footballers. Statathlon, 2017.

[SciSports, 2020] SciSports. SciSports and Sports Interactive will bring next-gen analytics to life in Football Manager. SciSports, 2020.

[UEFA, 2021] UEFA. UEFA Women's Champions League lays foundations for more sustainable future. UEFA, 2021.

[Worville, 2020] Tom Worville. Explained: How different are men's and women's football. The Athletic, 2020.