

East Coast Regional Datathon 2021

Team 24

March 2021

I. EXECUTIVE SUMMARY

IA. PROBLEM STATEMENT

In the years 2009-2012, Futbol Club Barcelona (Spain) came up with a unique strategy widely known as "Tiki-Taka"[1]. In this style of play, the entire team comes together in a cluster and maintain control of the ball in that cluster[2]. Following the Tiki-Taka strategy, Barcelona won 14 titles from 2008 to 2014. However, other club managers quickly devised counter-strategies to Tiki-Taka. These counter-strategies were so effective against Tiki-Taka that it practically died soon after 2012[3].

There does not exist a universal strategy that will always win the game. We expect, however, like most strategy games, there exist an assortment of viable strategies, each with its own winning and losing match-ups. As such, question we seek to answer is: do European football teams develop multiple meta-game strategies (metas)? If so, how do different metas interact with each other and how can we use these winning relationships to find the best strategy against a specific opponent?

IB. KEY FINDINGS

In this report, we present a *metagame analysis* on European football games between 2008-2016. Through clustering algorithms, we show there are indeed meta-game interactions between team strategies, each with its own winning and losing match-ups. From these clusters, we also identified various prominent features that could explain some of the match-up breakdown.

On the other hand, we have also identified one clusters as an anomaly as it is primarily defined by high overall ratings of players. This seem to suggest that the individual proficiency of players might be enough to offset strategic disadvantages in soccer. Furthermore, we were unable to identify 'cycles' of meta-strategies that dominate one another, but instead found multiple 'chains' of dominance. This seems to suggest that while there are meta-strategies in European football, other practical limitations in the game might make it difficult to see the full evolution of a game meta. In particular, we postulate that player contracts and inflexibility of teams in adopting new play styles could contribute to the lack of a fully developed metagame in European soccer.

Last but not least, using this meta-strategy table, we present a suggestion of counter strategies to adopt against other teams. Through a hypothesis testing procedure, we were able to reject the null hypothesis that "the best counter meta recommended by us does not increase the expected number of goals".

II. TECHNICAL EXPOSITION

IIA. DATA AND MEASURES

To minimize the effect of noise, our main goal in the data wrangling process is to create and identify all relevant attributes that could distinguish a particular meta-strategy from others.

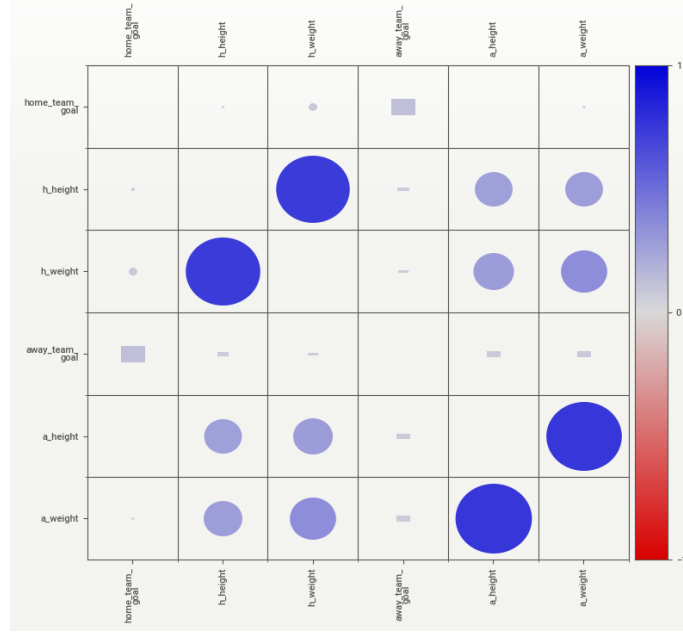


Figure 1: Correlation between height/weight information and home/away team goals

We attempted this by extending ‘match.csv’ with the team attributes as well as the aggregated player attributes for each team. The reason for including individual player attributes is so that we can contextualize the team attributes. For instance, we might expect a strategy of sending long balls to star strikers in the center and one of sending through balls to wing players to both have similar team attributes, but might have distinct breakdown in terms of preferred foot of players.

We have also made a decision to include the average player height and weight information for each team. While the dataset showed little correlation between the average height/weight and the goals scored, we identified that there is a possibility that height/weight information is only important in identifying particular meta-strategies but not goals scored in general. After all, it is well-known the advantages that height and build can bring to challenging the opposition in headers, possessions, and *set pieces*¹. We later found this to be an accurate assessment, as our some of the identified meta-strategies include height and weight attributes in its classifying decision tree.

Discrete attributes deserved their own unique treatments. We had to remove noisy discrete

¹A set piece refers to a play scenario when the ball is out of bounds, including corner kicks, free kicks and penalties

labels such as 'y' and 'stoc' under the 'attacking work rate' attribute. Upon reviewing the source of the data, we realized that there are only 3 true labels for this attribute, namely 'low', 'medium' and 'high'. Noisy labels are likely due to conversion errors from labels of other attributes such as the 'stocky' label under 'build' attribute which was excluded in the given dataset.

IIB. PRELIMINARY ANALYSIS

Our goal in the preliminary analysis is to see if our assumption on football metagames is valid. To this end, we run a simple K-means clustering algorithm on team-wide features that are processed from team_attributes.csv. We notice that many categorical columns in the data set are redundant information. Therefore, we take all numeric columns in team attributes to be features for K-means.

We normalize features by performing linear transformations on each feature so that the minimum and maximum values are transformed to 0 and 1, respectively. We choose min-max normalization over other normalization methods such as Normal standardization because the given data follow various unimodal distributions with similar minimum and maximum boundaries, as shown in Figure 2. Meanwhile, there are no outliers after our data cleaning process. This suggests robust scaling does not offer significant advantage either.

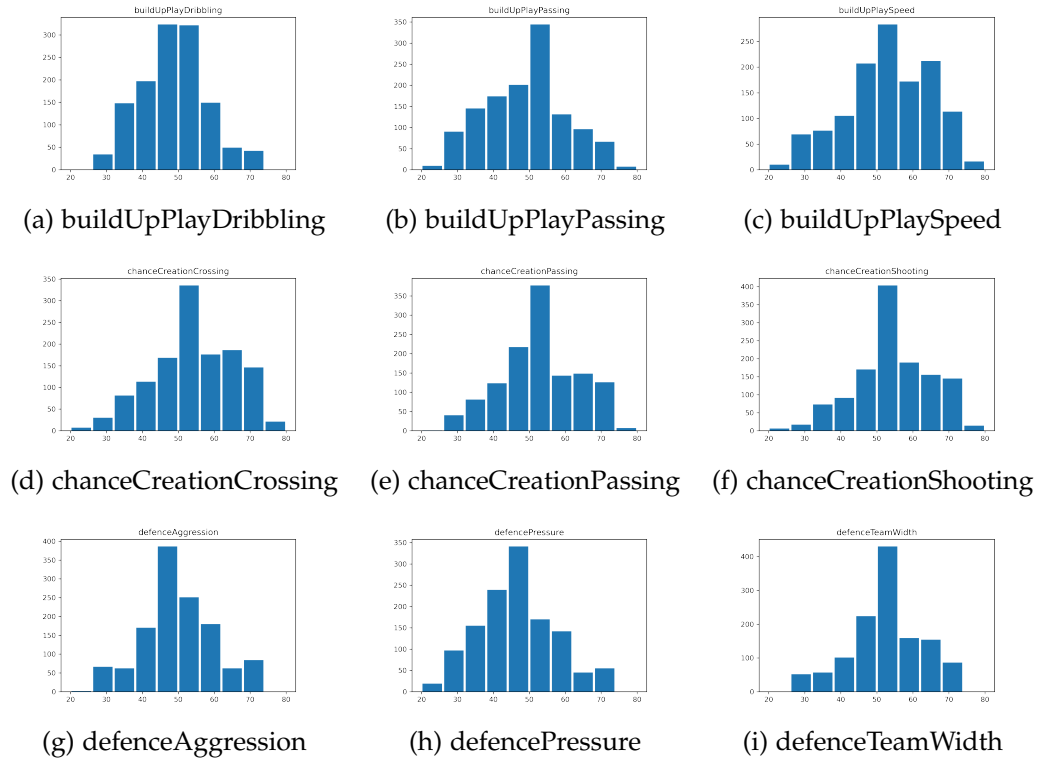


Figure 2: Distribution of each feature used in the preliminary K-means analysis

The number of clusters is a key variable in determining the usefulness and precision of the clustering algorithm. For our preliminary analysis, we use a simple "elbow curve" method to determine the appropriate number of clusters. The K-means algorithm is run on the data set with number of clusters varying from 1 to 50 and the overall fit score is examined, as shown in Figure 3(a). We see that the precision of clustering increased rapidly as the number of clusters is raised from 1 to 10, and continued to increase at a lower rate with more than 10 clusters. Here, 10 clusters are the point where a balance between precision and categorizing power is reached.

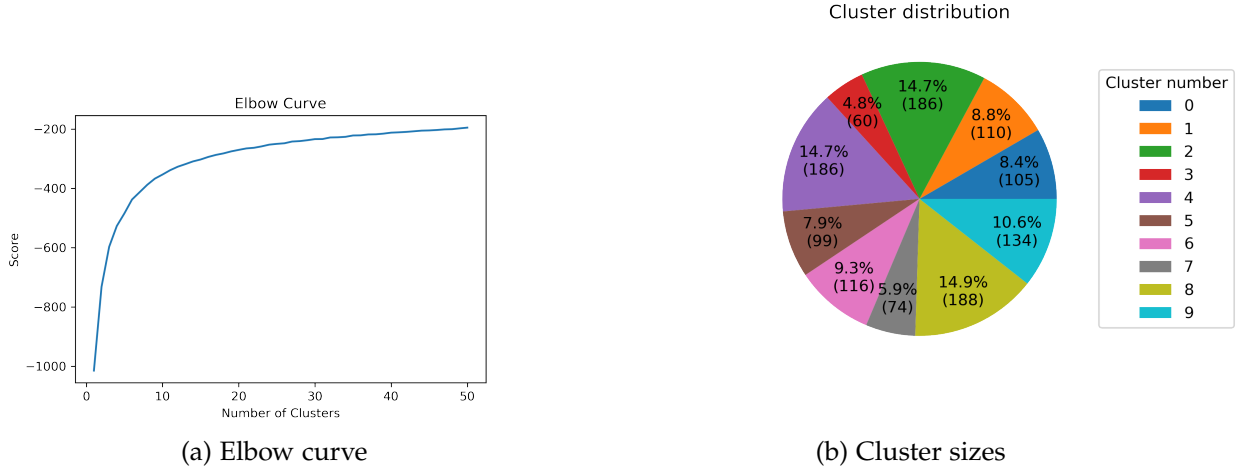


Figure 3: Elbow curve and cluster sizes in the preliminary K-means analysis

We note that the above clustering is performed on team attributes alone. Although no player attributes are taken into account, there is already a great diversity of team playstyles. The 10 clusters also have comparable sizes, which means the 10 metas we find here are all strategies that many teams can adapt to, as opposed to niche strategies that could only be played by one or two football players.

With this information, our assumption on teams having various metas is justified. That clustering gives poor precision (low score) when the number of clusters is lower than 10 is strong evidence that teams deploy different strategies and follow different playstyles.

IIC. CLUSTERING ANALYSIS

The K-means algorithm is a special case of clustering with a mixture of Gaussians where all variances are equal and mixture weights are equal. This assumption is clearly suboptimal for many real-life applications where the distribution of clusters are not spherical. Hence, with our objective to find the best meta strategy by clustering team play styles, we improved our clustering methodology with Gaussian Mixture Model.

Using the identical min-max normalized dataset with all team and players attributes, we first performed hyper-parameter search for GMM. There are two parameters of interest: number of clusters and the type of covariance matrix. To evaluate the performance of GMM with different

hyperparameters, we used penalized-likelihood criteria such as AIC and BIC. We chose both AIC and BIC together for model selection since a lower value of both AIC and BIC indicate a high data likelihood given the fitted GMM model and low overfitting.

By varying the number of clusters between 1 to 30 and the type of covariance matrix to be full, tied, and diagonal, we found that: 1) there is a large decrease in both AIC and BIC with 9 and 24 clusters 2) full covariance is computationally inefficient, and diagonal covariance type is best suited for meta strategy clustering (Fig 4).

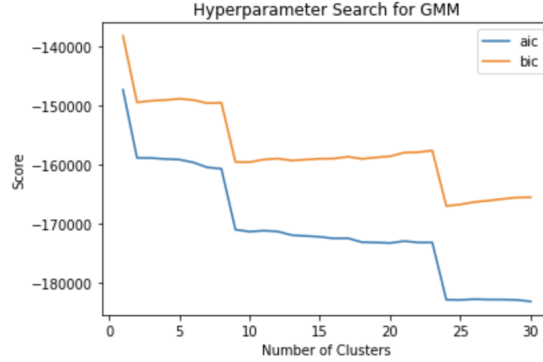


Figure 4: Finding the optimal number of clusters (9, 24) with diagonal covariance matrix in GMM, measured by AIC and BIC.

Hence, we formed 9 and 24 meta strategies and counted the number of team play styles in each meta strategy. We can see that 9 clusters provide a more general summary of meta strategy but 24 clusters provide a more detailed and balanced description of different meta strategies (Fig 5). We therefore decide to analyze meta interactions with clustering results of **both** 9 and 24 meta strategies.

Number of metas	Mean number of team play styles in meta	Variance of number of team play styles in meta
9	103	4269
24	38.8	544

Figure 5: Summary statistics for number of team play styles within each meta strategy after GMM clustering.

IID. INTER-META ANALYSIS

With the meta strategies from GMM, we want to understand and interpret the different meta strategies and how they interact with each other in football matches. We assigned each of the two teams at a match event to a meta strategy by looking for the closest clusters for the team’s play style. Since we treated different entries of the same team’s attributes as separate team instances, the same team (identified by team id) could still be assigned to different metas due to changes in playstyle or player attributes at different matches. This treatment allows for maximum accuracy in categorizing the playstyle of a team in a given point in time.

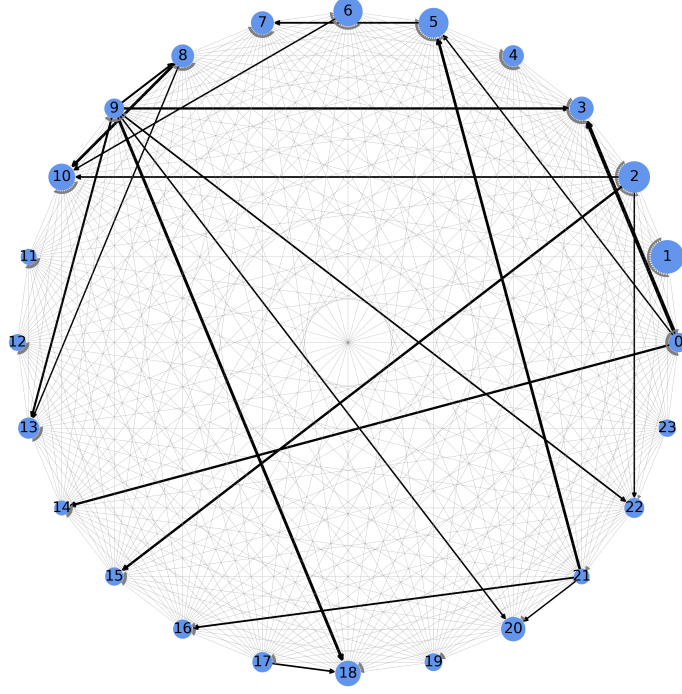


Figure 7: *Inter-meta winning relations between 24 GMM clusters.*

Despite the complex interaction between many pairs of the 24 meta strategies, we observe edges with significantly higher weights, indicating an even clearer patterns of inter-meta interactions and a large impact of meta strategy choice on the expected number of goals in a game. We also see some superior meta strategies, such as meta 9, that delivers a large number of goals against many other metas such as 10, 3, and 8. Fig 7 is also helpful as it points out interesting meta strategies with large impact on game outcomes for us to analyze later. Hence, we will proceed with only using the meta strategies from GMM with 24 clusters.

III. VALIDATION AND RECOMMENDATION TESTING

By analyzing inter-meta interaction, we gained a deeper understanding of how play styles that fall under one meta strategy are effective against play styles in some but not other meta strategies.

To verify results from inter-meta interaction and its value in real-life match strategizing, we seek to answer the following question: **given an opponent's meta strategy, does using the best counter meta from our meta interaction graph gives a statistically significant higher expected number of goals, than using a random meta strategy?** Here we define the best counter meta as the meta with the highest expected number of goals against the opponent's meta strategy in our meta interaction graph.

First, playing against a given meta strategy, we calculated the expected number of goals scored by an average counter meta strategy and the expected number of goals scored by the best counter meta strategy (Fig 8). For example, if our opponent is playing meta 1, a team using a random meta is expected to score 0.23 goals against our opponent. However, using the best counter meta suggested by our meta interaction graph, we should adopt meta strategy 9 and will have an expected number of goal of 1.36 against our opponent.

	meta	avg counter meta expected goals	best counter meta	best counter meta expected goals
0	0	-0.69	0	0.75
1	1	0.23	9	1.36
2	2	-0.52	0	1.03
3	3	0.64	0	4.00
4	4	0.20	0	1.39
5	5	0.25	21	3.00
6	6	-0.12	0	1.16
7	7	-0.11	5	2.00
8	8	-0.34	9	2.14
9	9	-0.82	9	0.44
10	10	0.39	8	3.00
11	11	0.12	0	1.36
12	12	0.02	12	0.32
13	13	0.24	9	2.31
14	14	0.11	0	2.50
15	15	0.32	2	2.62
16	16	0.06	21	2.00
17	17	-0.03	19	0.93
18	18	0.54	9	3.20
19	19	-0.14	5	0.50
20	20	0.10	9	1.68
21	21	-0.24	9	1.43
22	22	0.15	9	1.87
23	23	-0.00	7	0.60

Figure 8: *The expected number of goals by the best counter meta is significantly higher than that by an average counter meta against a given meta strategy*

To answer the question above, we perform a hypothesis testing by comparing the expected number of goals scored by an average counter meta and best counter meta. Although we have two samples (one for the average counter meta and one for our best counter meta), all observations are dependent to each other. Hence, we conduct a matched pair comparison by testing on the difference between the two samples. Assuming the number of goals is independent across different clusters and given a specific opponent meta, we wish to test

H_0 : The best counter meta recommended by our analysis
does not increase the expected number of goals.

versus

H_1 : The best counter meta recommended by our analysis
does increase the expected number of goals.

Mathematically, we test

$$H_0 : E_g^{BCM} - E_g^{AVG} = 0 \text{ versus } H_1 : E_g^{BCM} - E_g^{AVG} > 0$$

Since the sample size is small ($n=24<30$) and the population variance is unknown, we use t-statistic for hypothesis testing. Under the null hypothesis,

$$T = \frac{(E_g^{BCM} - E_g^{AVG})\sqrt{n}}{S} t_{n-1}$$

And we reject the null hypothesis if

$$T > t_{n-1, \alpha/2}$$

where S is sample variance, t_{n-1} is the Student's t-distribution with degree of freedom $n - 1$, and $\alpha = 0.05$. Plugging in the values in Fig 8, we indeed verify that the null hypothesis is rejected at the 0.05 significance level. There is statistically significant evidence that our analysis is capable of giving recommendations that increase the expected number of goals against a known opponent meta.

This result is highly non-trivial. Our model of selecting the best counter meta is immediately relevant to a team's manager: If they knew the meta strategy their opponent's team is playing, they could see if the best counter strategy is possible for their team. In cases where a certain strategy might not fit their players' abilities, the manager also has a sorted list of known metas that have higher chances of winning.

III. INTERPRETATION OF RESULTS

Throughout our analysis, we have 1) come up with 24 meta strategies with GMM, 2) studied inter-meta winning relations with weighted directional graph and 3) validated results using Best Counter Meta in increasing expected number of goals. The key question remains is to interpret what does meta strategy mean in the context of player and team attributes and how to form a meta strategy.

We designed a decision tree model to predict the meta strategy given a team's player style. We chose decision tree because it highly interpretable, and we used the structure of the decision tree to generate rules for meta strategy assignment.

class_name	instance_count	rule_list
0	11	Num:[3.000] (gk_kicking <= 22.170) and (defenceDefenderLineClass_Cover > 0.500) and (reactions > 69.111) and (buildUpPlayDribbling > 40.750) and (long_shots > 60.619) and (buildUpPlayDribbling <= 46.250)
1	70	Num:[8.000] (gk_kicking <= 22.170) and (defenceDefenderLineClass_Cover > 0.500) and (reactions <= 69.111) and (volleys <= 42.240) and (reactions > 63.576) and (finishing <= 45.219) Num:[24.000] (gk_kicking <= 22.170) and (defenceDefenderLineClass_Cover > 0.500) and (reactions <= 69.111) and (volleys > 42.240) and (shot_power <= 58.297) and (short_passing > 60.942)
2	57	Num:[43.000] (gk_kicking <= 22.170) and (defenceDefenderLineClass_Cover > 0.500) and (reactions <= 69.111) and (buildUpPlayDribbling > 40.750) and (long_shots > 60.619) and (buildUpPlayDribbling > 46.250)
3	41	Num:[2.000] (gk_kicking <= 22.170) and (defenceDefenderLineClass_Cover > 0.500) and (reactions <= 69.111) and (volleys <= 42.240) and (reactions > 63.576) and (finishing <= 45.219)

Figure 9: Generating rules for meta strategy assignment

Here we refer to Fig 8 on the expected number of goals by the best counter meta.

Looking at the summary, a few features stand out. First of all, meta-strategy 9 leads all strategies in terms of average goals against other strategies, as well as having only itself as its most significant counter. On the other hand, we also witness chains of ‘dominance’ such as

$$21 \rightarrow 5 \rightarrow 19 \rightarrow 17 \rightarrow 7 \rightarrow 23$$

where the preceding strategy trumping the proceeding ones in the chain above. While our analysis only provides correlation and hence is limited in its ability to explain the causal factors behind such observations, we can make an attempt based on the learned categorization rules of our meta-strategies.

In particular, we noted after running a decision tree categorization on the meta-strategies that strategy 9 leads all strategies in terms of the average overall rating and potential rating of its players. On the other hand, however, strategy 9 has average to low scores for team attributes such as dribbling and passing. This strategy, therefore, almost seem to present itself as an antithesis to the hypothesis that there exist meta-game strategies in soccer. We will look further into this issue in a later section.

On the other hand, the identified chain of dominance is particularly interesting as we can identify defining traits of each meta-strategy that might explain some of the relationships.

For instance, upon running a simple decision tree algorithm on the identified meta-strategies, we saw that strategy 21 can be defined by the high proportion of teams that adopt an *offside trap* defensive maneuver. In soccer, the offside trap works by having a defensive line far up on the field, such that opposing strikers run the risk of getting fouled by rule of *offside*.²

The offside trap is particularly effective against opponents who make medium to long range passes, which are then susceptible to having receivers be in an offside position. On the other hand, however, it can be countered by build up plays that heavily rely on dribbling. Incidentally, we saw that meta-strategy 5, which 21 dominates, can be identified by teams with *Buildup Play Dribbling*

²In soccer, an offside foul is when an attacker is beyond the second last enemy player before a ball is passed to him

of < 40.8 and *Buildup Play Passing* of > 46.5 . This suggests that teams which rely on passing up the field tend to suffer against teams of meta-strategy 21, just as our intuitive analysis indicates.

Interestingly, however, we did not manage to identify any obvious (high deviation in win rate from 50-50) ‘cycles’ of dominance. While chains of dominance do exist in most strategic games, in developed metagames it is often the case that only cycles survive eventually. For instance, a cycle would be a set of three meta-strategies that follow a ‘scissor-paper-stone’ relationship.

This is because in a strategic game, once identified, no team will adopt the lowest lying meta-strategy in a chain of dominance. This will in turn cause the second-lowest meta-strategy to be the weakest. Inductively, no team will adopt any strategy other than the top-level meta-strategy. Where there is a cycle, however, we will see a percentage of teams adopting every possible meta-strategy in the cycle.

We unfortunately were unable to identify clear cycles in our meta-strategy clusters. However, we believe that this could be due to a structural limitation of soccer in the first place. Traditionally, in a strategic game, teams would change their play styles to adopt the most effective meta-strategies available. This assumes however that this change of play styles is quick to implement and will be as effective as the masters of the particular play style. This is simply not the case in soccer. Soccer teams are limited by player contracts, so they cannot switch out players with ease. They are also limited by the fact that it takes time, team cohesion and coaching knowledge to adopt new play styles on the fly. As a result, the soccer meta did not evolve as expected.

III. LIMITATIONS

Due to limited time and resources, we are aware of various weaknesses of our approach.

In calculating player attributes, we took the aggregated mean over matches for each team. While this is a convenient way of integrating player attributes into each team, this ignores the fact that different features have different impacts on distinct roles. For instance, defensive ratings might not apply as much for strikers, while offensive ratings might not be as applicable for defenders. Goalkeeper ratings, on the other hand, are almost irrelevant to field players.

Furthermore, the biggest question mark on our results is the existence of meta-strategy 9. This is the most dominant strategy cluster, with only itself as the sole reliable counter, but it turns out from looking deeper that the great identifying feature is the pure over skill rating of the team’s players. Rather than disproving our ‘meta-game’ hypothesis of football leagues, however, it is not unreasonable to expect that with sufficiently strong players, certain strategic disadvantages can be neutralized. Furthermore, attribute information might suffer from bias as they could be affected by team play style in the first place. Indeed, teams with an innately aggressive and fast play style could allow players to demonstrate higher levels of acceleration, for example. Teams that relied on set pieces could have strategies that relied on mid-ranged passing and drawing opposition attention. These are examples of hidden attributes that produce the given attributes as a consequence, instead of the other way round. Our models make use of the given attributes as a starting point, and as such might be limited in identify these hidden structures.

III. CONCLUSION

In conclusion, through the datasets we have shown that we can identify distinct meta-strategies. As validation, we identified certain characteristics in our meta-strategy table that has reasonable real-life justifications. We did make various simplifying assumptions that might have limited the usefulness of our results; for instance, the fact that a meta-strategy that simply identify teams with 'the best players' in meta-strategy 9 does not seem useful in identifying team playstyles.

With more time and resources, we see great potential in removing such anomalies. We could possibly model each meta-strategy's win rate as a base rate scaled by the average rating of the team's players. This will take into account the fact that a team with stronger players might be able to neutralize meta-strategic disadvantages.

Furthermore, we mentioned under limitations that the given attributes could be secondary features that arise from hidden attributes of meta-game strategies. To explore this issue, we could make use of neural nets in a manner which allows us to identify hidden structures behind the player attributes.

REFERENCES

- [1] "The Evolution of Barcelona's Tiki Taka." The Evolution of Barcelona's Tiki Taka | Goal.com, www.goal.com/en/news/12/spanish-football/2015/09/29/15804882/the-evolution-of-barcelonas-tiki-taka.
- [2] "The Evolution of Barcelona's Tiki Taka." The Evolution of Barcelona's Tiki Taka | Goal.com, www.goal.com/en/news/12/spanish-football/2015/09/29/15804882/the-evolution-of-barcelonas-tiki-taka.
- [3] Desmond, Rhys. "Counter Attacking and the Death of Tiki-Taka Football." The Mastermind-Site, 7 Feb. 2021, themastermindsite.com/2019/09/21/counter-attacking-and-the-death-of-tiki-taka-football/.