IJCCS (Indonesian Journal of Computing and Cybernetics Systems)
Vol.x, No.x, July xxxx, pp. 1~5
ISSN (print): 1978-1520, ISSN (online): 2460-7258
DOI: 10.22146/ijccs.xxxx ■
1

World Cup 2022 Knockout Stage Prediction Using Poisson Distribution Model

Stanislaus Jiwandana Pinasthika[1], **Dzikri Rahadian Fudholi*[2]**
[1,2]Department of Computer Science and Electronics, FMIPA UGM, Yogyakarta, Indonesia
e-mail: [1]stanislausjp@gmail.com, *[2]**dzikri.r.f@ugm.ac.id**

Abstrak

Sepak bola merupakan salah satu olahraga terpopuler di dunia. Setiap topik yang berkaitan dengan sepakbola menarik untuk dibahas termasuk prediksi pemenang pertandingan Piala Dunia FIFA. Topik ini tidak hanya sebagai topik diskusi santai namun, dapat menjadi penunjang pengambilan keputusan bagi jajaran kepelatihan dalam menilai kesiapan tim. Beberapa prediksi menggunakan dataset pertandingan yang sangat besar namun hal tersebut tidak relevan karena setiap edisi piala dunia, tim nasional akan memiliki komposisi tim yang berbeda. Sehingga, dibutuhkan prediksi pertandingan dengan data yang lebih relevan. Model distribusi Poisson digunakan untuk prediksi pertandingan babak fase gugur di Piala Dunia FIFA 2022. Probabilitas menang dan kalah dihitung menggunakan rata-rata gol yang dicetak dan gol yang dicetak lawan, lalu dievaluasi perbedaan dengan hasil aktualnya menggunakan jarak de Finetti. Dari 15 pertandingan di fase gugur, ada 8 prediksi yang dapat diterima dengan 6 diantaranya adalah pertandingan di 16-besar. Selain menerapkan Poisson, penelitian ini juga untuk menunjukkan bahwa dataset yang terbatas masih bisa memecahkan masalah prediksi. Untuk penelitian selanjutnya dibutuhkan atribut data baru untuk membentuk lambda Poisson. Untuk meningkatkan ketepatan prediksi, perlu ditambah 3-4 data pertandingan jelang Piala Dunia.

Kata kunci*—prediksi probabilistik, jarak de Finetti, prediksi pertandingan sepak bola*

Abstract

Football is one of the most popular sports in the world. The popularity makes every topic related to football interesting, for instance, the FIFA World Cup winner prediction. This topic is not only for casual discussion but could be a practical decision support for coaching staff to rate the team's readiness. Most prediction methods use large match datasets. Since every national team has a different squad for every world cup and the FIFA World Cup is held every four years, the usage of a large match dataset is irrelevant. Therefore, there is a need for a prediction method based on the relevant data. We applied the Poisson distribution model for predicting the FIFA World Cup 2022 knockout stage match results. We calculate the probability of winning and losing based on their average goal scores and goal conceded and evaluate the difference by the actual result using de Finetti distance. The successful prediction is 8 out of 15 matches, with six inside the round of 16 games. This prediction model is also a brief example to overcome prediction problem using limited dataset. Thus, the new data attributes need to reformulate Poisson's lambda. Further studies need to add the 3-4 prior world cup matches data to increase the acceptance of prediction.

Keywords—*probabilistic prediction, de Finetti distance, football match prediction*

Received June 1st,2012; Revised June 25th, 2012; Accepted July 10th, 2012

1. INTRODUCTION

Football is one of the most popular sports in the world. The popularity of football has a good influence on economic development [1]. The annual football league always has its audiences and earns economic benefits from them. Selling merchandise and recruiting world-class players are several things that can be the profit resource. FIFA, a global football organization, governs a prestigious football tournament called FIFA World Cup that increases the nation's host welfare [2]. For instance, the FIFA World Cup 2018 has had a cumulative effect on the Russian economy, amounting to USD 13 billion, representing roughly 1% annual GDP between 2013 and 2018.

Since the popularity and economic profits, every discussion topic related to FIFA World Cup is interesting. One of the most popular topics is the FIFA World Cup winner prediction discussed by football fans and sports journalists [3]. Football prediction also helps the coaching staff to rate the team's readiness [4]. An accurate winner prediction with a high winning probability will be good decision support for the coaching staff to set the starting eleven for the next match.

Several methods are applied to solve the prediction problem based on a probabilistic model or machine learning approach. Some of them are Poisson models [5], an attention-based LSTM network [1], a hybrid of LSTM and RNN [4], a combination of ANN and DNN [6], and a Priori prediction algorithm [7]. Most prediction models train many features of team performance data, such as the goals, ball possession, corners, crosses, or pass accuracy, which is necessary winning factors. Previous researchers usually require large datasets to obtain the optimal prediction model and acquire the match data as much as possible from the earlier seasons that they could discover to the latest one.

Occasionally, a massive amount of data is not always representing a good accuracy prediction. The irrelevancy of the majority causes this occasion of used data. For instance, obsolete and aged data may not influence data nowadays. The limited data resource also leads to an inadequate prediction model. This case could be overcome by augmenting the current data. But the inappropriate augmented method for adding the current data may lead to a decline in the accuracy [8].

An accurate prediction model is constantly supported by valid data with pertinent attributes [9]. Even using the correct data does not guarantee the production of a precise prediction model and often meets uncertainty [10]. Hence, the relevant data source needs to be sought. To predict a football match's result, match-related data, such as the number of goals, ball possession, number of shots, etc., are needed. The main factor of the mentioned attributes is the players. Unlike the regular football league, the FIFA World Cup is held annually every four years. The national team used to call up specific players that might be different from the friendly match or other tournament player calls.

Particularly, the result of the previous edition of the World Cup is irrelevant, so the only data which is perfectly fit in this study are the data group stage results from the current edition of World Cup. Additionally, the football match is also such a rare event. Every team has a fair win, draw, or lose probability. A team could score many goals in the previous games and still possibly lose in the next games. Hence, the Poisson distribution model is proposed to predict the winner of the FIFA World Cup 2022. The Poisson distribution model could be used to develop a prediction model in that the probability of being repeated is rare [11]. Furthermore, the number of goals is always an integer that is suitable for the Poisson model to predict the winner of a football match [12].

Besides giving the prediction to the World Cup knockout stage matches, the applied Poisson distribution model is also expected to be a solution to overcome prediction problem with limited dataset. This study offers a brief match prediction solution to all coaching staff, not limited to football, during the competition. Application of the Poisson distribution model is expected being a brief example to overcome prediction problem using limited dataset.

2. METHODS

This study uses the Poisson distribution model to predict the winner of the match. This prediction uses the Poisson distribution since the football match problem has fixed time intervals, and the observed number of events is random [13].

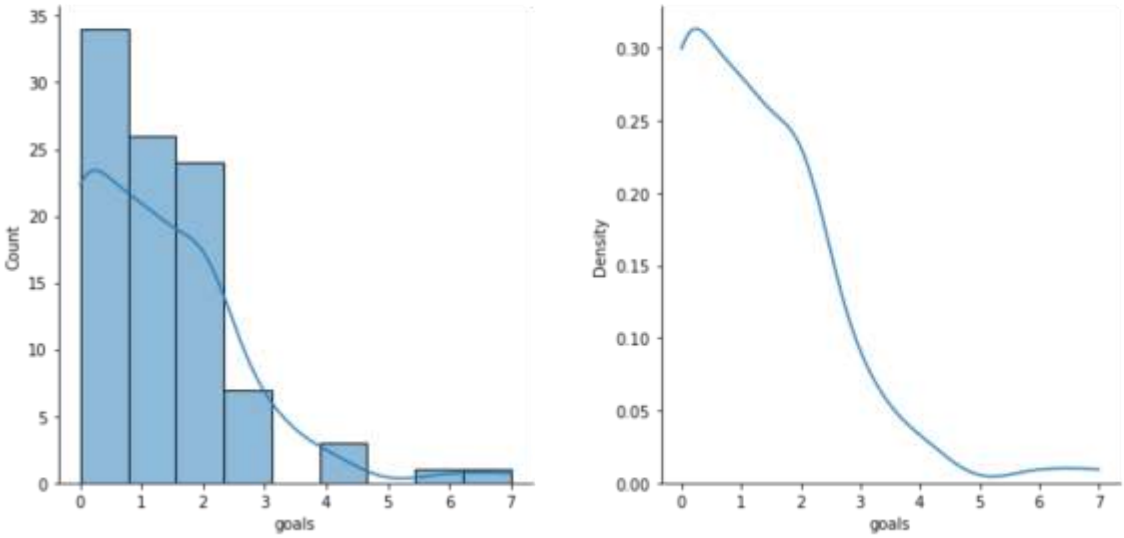

Figure 1: The 2022 FIFA World Cup group stage goal distribution

The number of goals scored and conceded is a discrete probability distribution. A team has more chance to win the match if they have better offensive and defensive aspects, as shown by the number of goals they scored more than the goal they conceded [14]. Individual match probability does not affect the likelihood of another match; thus, the Poisson method can be used.

Figure 1 shows the distribution of goals during the group stage of the FIFA World Cup 2022. Both the left and right figures explain how many goals have been created by a team in a match. The left figure showed the number of goals, while the right displayed the densities. The higher figure was the team that ended the game without scoring a goal. Scoring a goal and two goals is the second-highest and third-highest figure, respectively. However, each figure is independent and called discrete distribution.

2.1 Poisson Distribution

A game in the knockout stage The FIFA World Cup has results win for the home team or away team without a home and away format, so both teams have an equal chance to win the game. All teams in the knockout stage have already played three matches. The average goal they created represents the offensive strength, and the defensive aspect is represented by the average conceded. A team has a higher goal-created chance if they have a higher average of goals scored, and the opponent has fewer chances to keep their goalpost from conceding. The strength of each team was represented by lambda (λ), as shown in equation (1).

Two lambda values were needed to determine the strong point of teams *a* and *b* winning the match *m*. Each lambda was included in the Poisson formula in equation (2). The Poisson formula calculated the scoring probability of each team. Iteration *x* represented possible goals created by each team in a single match. Assume the knockout stage is more competitive than the group stage, and all the contestants have similar performances. The maximum number of goals scored in the group stage is 7. Thus, the maximum *x* was set to 7.

$$\lambda_a = goalscored_a \times goalconceded_b, \lambda_b = \ goalscored_b \times \ goalconceded_a \tag{1}$$

$$f(x;\lambda) = P(X = x) = \frac{(e^{-\lambda}\lambda^{-x})}{(x!)}, x = 0,1, \dots, 7 \tag{2}$$

$$p(x_a, x_b) = P(X_a = x_a) \times P(X_b = x_b) \tag{3}$$

After the scoring goals probability of each team was discovered, we needed to simulate all possible results that could happen at the end of the match. Hence, we multiplied each other the scoring goals probability, as shown in equation (3). The multiplication was grouped by the results, such as a win, b win, or both teams were drawn.

The pseudo-code of predicting process is in Algorithm 1, where lambda a and lambda b are the input of this process for obtaining the winning possibility of both teams. The primary procedure was written in line 4 based on equation (3), where the Poisson formula of both teams is multiplied. The multiplication Poisson was grouped to the prepared variables based on the results as the probability of the match could be ended by winning team a, team b, or a draw. Hence, the likelihood of each team winning the game was written in lines 16 and 17.

Algorithm 1: Predicting Winner Chance
Input : λ_a, λ_b
Output : points_a, points_b

```
1:  prob_a, prob_b, prob_draw = 0,0,0
2:  for x_a in [0,1,2,3,4,5,6,7] do
3:      for x_b in [0,1,2,3,4,5,6,7] do
4:          p_total = p( x_a, λa) * p( x_b, λb)
5:          if x_a == x_b then
6:              prob_draw += p_total
7:          end if
8:          else if x_a > x_b then
9:              prob_a += p_total
10:         end if
11:         else if x_a < x_b then
12:             prob_b += p_total
13:         end if
14:     end for
15: end for
16: points_a = 3 * prob_a * prob_draw
17: points_b = 3 * prob_b * prob_draw
```

Each point was obtained by multiplying the winning probability by three, the number of points given to the match’s winner. The draw probability was included in both team points since the knockout stage match requires the winner to go to the next round.

2. 2 De Finetti measure

We use the de Finetti distance for measuring the prediction distance to the actual result. Every joint probability distribution over possible sequences of an event is given a suitable latent independent variable [15]. The development of the football match is independent and subjective.

As a result, not mandatory close to the frequency, as de Finetti argues from the subjectivist point of view.

All probable match results are included in the simplex set, as shown in equation (4), with P_a the winning chance of team *a* and P_b the winning probability of team *b*. Both winning chances P_a, P_b are represented by (1,0) and (0,1), respectively.

$$S = (P_a, P_b) \epsilon \mathbb{R}^2 : P_a + P_b = 1, P_a \geq 0, P_b \geq 0 \quad (4)$$

The de Finetti distance formula forms like the not-rooted Euclidean distance. The distance between the prediction and the actual result is given by equation (5), where b_1 and b_2 are represented by (1,0) and (0,1), respectively. The distance is acceptable if the space is under the threshold. The de Finetti threshold is the distance when both chances are equivalent. Assume the $P_a + P_b$ is 1 and team *a* win over team *b*, then the de Finetti threshold is *th* = $(0.5 - 1)^2 + (0.5 - 0)^2 = 0.5$.

$$d = (P_a - b_1)^2 + (P_b - b_2)^2 \quad (5)$$

For instance, if the Poisson predicts that team *a* will win over team *b* and the winning chance is (0.560621, 0.188412) with the team *a* winning the match, then the de Finetti distance is *d* = $(0.560621 - 1)^2 + (0.188412 - 0)^2 = 0.228553$. Subsequently, the prediction is acceptable.

3. RESULTS AND DISCUSSION

The proposed method to predict The FIFA World Cup 2022 knockout phase winner was using the Poisson method and group stage match data to obtain average goals created and goals conceded for each team. Goals scored and goals conceded are the attributes with the most obvious influence on the team's victory. Every football tactic is created for scoring more goals against the opponent and avoiding their own goal being conceded. Our data source is Google after match data which contains several attributes like the number of shots, number of shots on targets, passing accuracy, number of offsides, number of fouls, and number of corners. We reformat the obtained data and add several attributes, depending on the team, such as the result labels. The reformatted data is shown in Figure 2.

The match data, as shown in Figure 2, was sorted by each team. This format effectively depicts team performance and the result they got. Every team in the world cup did their best to win the match, such as winning the possession, passing the ball toe-by-toe, maintaining passing accuracy, or shooting to the target more often than the opponent. Those attributes are not always proportional to the result. For instance, Argentina versus Saudi Arabia, which is depicted in indexes 8 and 9, was won by Saudi Arabia. However, the amount of passing that Saudi Arabia made is less than Argentina made. Argentina attempted to make more shots than Saudi Arabia, but at the final whistle, Saudi defeated Argentina. This match led us to conclude that the only attribute that positively correlates with results is the goals.

	date	nationalTeam	matches	goals	penalty_shootout	shot	shot_on	posession	pass	acc_pass	fouls	y_card	r_card	offside	corner	result
0	20/11/2022	Qatar	gs1	0	0	5	0	0.47	434	0.80	15	4	0	3	1	L
1	20/11/2022	Ecuador	gs1	2	0	6	3	0.53	486	0.84	15	2	0	4	3	W
2	21/11/2022	England	gs1	6	0	13	7	0.79	797	0.89	9	0	0	2	8	W
3	21/11/2022	Iran	gs1	2	0	8	3	0.21	215	0.66	14	2	0	2	0	L
4	21/11/2022	Senegal	gs1	0	0	15	4	0.46	385	0.79	13	2	0	2	6	L
5	21/11/2022	Netherland	gs1	2	0	10	3	0.54	436	0.81	13	1	0	1	7	W
6	22/11/2022	USA	gs1	1	0	6	1	0.59	567	0.87	15	4	0	1	5	D
7	22/11/2022	Wales	gs1	1	0	7	3	0.41	403	0.76	10	2	0	1	3	D
8	22/11/2022	Argentina	gs1	1	0	15	6	0.70	596	0.85	7	0	0	10	9	L
9	22/11/2022	Saudi Arabia	gs1	2	0	3	2	0.30	264	0.67	21	6	0	1	2	W

Figure 2 The overview of FIFA World Cup 2022 group stage data

Since the purpose of this study is to predict the result of the knockout stage match, the data of the team which did not make it through are unselected. Subsequently, the average goals scored and the average goals conceded by each knockout stage contestant need to be obtained, as shown in Table 1. The obtained average data are from three group stage matches, which each contestant has done. The average goals scored and the average goals conceded data represented the attacking and defensive strength, respectively.

Almost all knockout stage contestants score more goals than their opponents, meaning they have proper attacking strength for going through the next stage. England and Spain have higher average goals scored after making nine (9) goals in total at the group stage. Otherwise, a team has excellent defensive strength if a relatively low average is conceded. Several contestants have an effective defending system with just conceded 0.333 goals per match, such as the USA, Brazil, Morroco, Croatia, and the Netherlands.

Table 1 Average goals scored and goals conceded by the knockout stage contestants

National Team	Average Goals Scored	Average Goals Conceded
Argentina	1.667	0.667
Australia	1.000	1.333
Brazil	1.000	0.333
Croatia	1.333	0.333
England	3.000	0.667
France	2.000	1.000
Japan	1.333	1.000
Morocco	1.333	0.333
Netherland	1.667	0.333
Poland	0.667	0.667
Portugal	2.000	1.333
Rep. Korea	1.333	1.333
Senegal	1.667	1.333
Spain	3.000	1.000
Switzerland	1.333	1.000
USA	0.667	0.333

Table 1 is needed to find out the lambda and calculate the Poisson. For instance, the group A winner, the Netherlands, had to face the runner-up of group B, the USA. The lambda for the Netherlands is their own average goals scored, 1.667, times the average goals conceded of the

USA, 0.333, and vice versa for the lambda of the USA. Subsequently, the Poisson points for the possibility of each goal can be obtained.

Netherlands and USA have the average goals scored better than their average goals conceded. However, the average number of goals scored by the Netherlands is better than the average number of goals scored by the USA. Both teams have a similar average of goals conceded. Table 2 shows the Poisson points for every possible result. Since the Netherlands team scored better average goals, their chance to win the game was higher. Moreover, Poisson predicted that the match might be ended up with either team winning the game with a three (3) or more goals difference.

Table 2 Poisson points for every possible results

Netherland \ USA	0	1	2	3	4	5	6	7
0	0.459	0.102	0.011	0.0008	4.66e-05	2.07e-06	7.67e-08	2.43e-09
1	0.255	0.057	0.006	0.0005	2.59e-05	1.15e-06	4.26e-08	1.35e-09
2	0.071	0.016	0.0017	0.0001	7.18e-06	3.19e-07	1.18e-08	3.75e-10
3	0.013	0.003	0.0003	2.39e-05	1.33e-06	5.90e-08	2.19e-09	6.93e-11
4	0.002	0.0004	4.49e-05	3.32e-06	1.84e-07	8.19e-09	3.03e-10	9.62e-12
5	0.0002	4.48e-05	4.98e-06	3.69e-07	2.05e-08	9.09e-10	3.37e-11	1.07e-12
6	1.87e-05	4.15e-06	4.60e-07	3.41e-08	1.89e-09	8.42e-11	3.12e-12	9.88e-14
7	1.48e-06	3.29e-07	3.65e-08	2.71e-09	1.50e-10	6.67e-12	2.47e-13	7.83e-15

Table 3 Acceptance of the knockout stage match winner prediction with the group stage data only

Team A	Team B	Match	Team A Wins	Team B Wins	Actual Winner	De Finetti distance	Acceptable
Netherland	USA	Round of 16	0.560621	0.188412	Netherland	0.228553	Yes
Argentina	Australia	Round of 16	0.376829	0.050455	Argentina	0.390888	Yes
France	Poland	Round of 16	0.451508	0.159615	France	0.326320	Yes
England	Senegal	Round of 16	0.185231	0.013783	England	0.664038	No
Japan	Croatia	Round of 16	0.105282	0.50567	Croatia	0.255446	Yes
Brazil	Rep. Korea	Round of 16	0.50567	0.105282	Brazil	0.255446	Yes
Morocco	Spain	Round of 16	0.367052	0.232383	Morocco	0.454625	Yes
Portugal	Switzerland	Round of 16	0.276031	0.221401	Portugal	0.573150	No
Croatia	Brazil	Quarter Final	0.437559	0.30973	Croatia	0.412273	Yes
Netherland	Argentina	Quarter Final	0.478748	0.178948	Argentina	0.903326	No

Morocco	Portugal	Quarter Final	0.430003	0.090535	Morocco	0.333093	Yes
France	England	Quarter Final	0.062349	0.301492	France	0.970087	No
Argentina	Croatia	Semifinal	0.240071	0.457257	Argentina	0.786576	No
France	Morocco	Semifinal	0.159615	0.451508	France	0.910106	No
Argentina	France	Final	0.324320	0.220361	Argentina	0.505102	No

All the Poisson points in Table 2 are accumulated according to the predicted score as shown in the Team A Wins, and Team B Wins columns part of Table 3. The got winning chance of the Netherlands is 0.560621, and the USA has a winning chance amount of 0.188412. Therefore, we concluded that the Netherlands was predicted to go to the quarter-final by Poisson distance. The actual winner of the match was the Netherlands, but the prediction absolutely could not consider a good prediction. We need to obtain the acceptability of the prediction using de Finetti distance. The de Finetti distance measured the distance between the actual and predicted results. The smaller the de Finetti value is, the closer prediction has been made to the actual result. We set the threshold of de Finetti acceptance in the amount of 0.5 since the knockout stage has only two outcomes, win or lose. In Table 3, the acceptance of predictions will be labeled as "Yes" if de Finetti's measured distance is lower than the threshold, as mentioned in subchapter 2.2.

The Poisson distribution model was already predicting the winner flawlessly in the round of 16. However, two of eight predictions of a round of 16 are unacceptable due to more than the de Finetti threshold. Several predictions did not pass the threshold slightly—for instance, the match between Portugal and Switzerland in the round of 16. Portugal was already winning the game, as predicted by the Poisson distance. The difference in the winning chances for Portugal and Switzerland is slight, so it is difficult to find the de Finetti distance. Since we put the draw point for both teams, the match result could be predicted as a draw.

The acceptance of the predictions declined dramatically. The more advance the knockout round is, the more unpredicted the match-winner prediction is. The number of unaccepted predictions is nearly half of the total 15 matches. As each candidate wants to show their top performance in the knockout and increase their chance to win the World Cup, using just group stage average goals data is obsolete. This problem led to appending the latest match data and updating the average goals data.

Table 4 Acceptance of the knockout stage match winner prediction with additional data

Team A	Team B	Match	Team A Wins	Team B Wins	Actual Winner	De Finetti distance	Acceptable
Croatia	Brazil	Quarter Final	0.319472	0.355382	Croatia	0.589415	No
Netherland	Argentina	Quarter Final	0.406344	0.172256	Argentina	0.850276	No
Morocco	Portugal	Quarter Final	0.424632	0.112504	Morocco	0.343705	Yes
France	England	Quarter Final	0.045823	0.297482	France	0.998949	No
Argentina	Croatia	Semifinal	0.27759	0.329576	Argentina	0.630497	No
France	Morocco	Semifinal	0.127129	0.475847	France	0.988334	No
Argentina	France	Final	0.190061	0.274212	Argentina	0.731193	No

We recalculated the winning chance of each team for the quarter-final, semifinal, and final match using the updated average goals data. We put the latest match data before predicting a match in a particular phase. For instance, we averaged the goals scored in the group stage until the semifinal as the substance of Poisson lambda in the final match prediction and the goals from the group stage until the quarter-final for predicting the semifinal match.

In the lambda formula, we included the penalty-shootout goals created and conceded. The weight of the penalty-shootout goal is considered half of the average goal. The goal created in the full-time is the pure result of the game plan. Eleven men were involved in the process of goal creation. The penalty given in normal time has full weight because it is caused by a foul in the penalty area, in which the attacking side's movement and the defending side's effort already exist. Otherwise, the penalty shootout is a mind game between the goalkeeper and the executor and is committed after both teams have equal numbers of goals in the normal time.

Table 4 summarizes the result of recalculating the winning prediction from the quarter-final to the final. This method of averaging goals is unacceptable, and the matches seem hard to predict. Six of seven match predictions differ from actual matches, and the highest de Finetti distance is in the quarter-final match France against England, amount 0.998949. So, the latest match result should not be added to the average goals scored and conceded as the Poisson lambda.

4. CONCLUSIONS

This study used the Poisson distribution model to predict the FIFA World Cup's knockout stage result. We use the average of goals scored and the average of the goal conceded data as the Poisson lambda. We observed the utilization of only group stage data, outcoming better prediction results than adding the group stage data with the latest match data before predicting the current stage match with 6 of 8, the round of 16 matches, and 2 of 4, the quarter-final match is correctly predicted. The complexity of winning factors in the knockout stage matches makes the goal attribute inappropriate to represent the whole team's performance.

The Poisson distribution model is still useable as a prediction model as long as the lambda is composed of proper and related data. We suggest generating new football data attributes such as expected goals (xG), saves percentage, shot on target, etc., for increasing the acceptance of football match result prediction. Adding 3-4 of the pre-world cup matches data is also recommended.

REFERENCES

[1] Q. Zhang, X. Zhang, H. Hu, C. Li, Y. Lin, and R. Ma, "Sport match prediction model for training and exercise using attention-based LSTM network," *Digit. Commun. Netw.*, vol. 8, pp. 508–515, Jan. 2021, doi: https://doi.org/10.1016/j.dcan.2021.08.008.

[2] S.Subathra *et al.*, "Economic Implications of Hosting FIFA World Cup – A Study with Social Reference to South Africa, Brazil, Russia, and Qatar," in *FIFA 2022 Qatar, The Legacy,* Nova Science Publishers, Inc., 2021.

[3] A. A. Azeman, A. Mustapha, N. Razali, Nanthaamornphong, and Wahab, Mohd Helmy Abd, "Prediction of Football Matches Results: Decision Forest Against Neural Networks," *Int. Conf. Eletr. Eng. Comput. Telecommun. Inf. Technol.*, vol. 8.

[4] E. Tiwari, P. Sardar, and S. Jain, "Football Match Result Prediction Using Neural Network and Deep Learning," *Int. Conf. Reliab. Infocom Technol. Optim. Trends Future Dir.*, vol. 8, Jun. 2020.

[5] S. N. Maozad, S. N. A. M. Razali, A. Mustapha, A. Nanthaamornphong, M. H. A. Wahab, and N. Razali, "Comparative Analysis for Predicting Football Match Outcomes based on Poisson Models," *Int. Conf. Eletr. Eng. Comput. Telecommun. Inf. Technol.*, vol. 19, 2022.

[6] Md. Ashiqur Rahman, “A deep learning framework for football match prediction,” *Springer Nat. Appl. Sci.*, Jan. 2020, doi: https://doi.org/10.1007/s42452-019-1821-5.

[7] B. Siswanto, M. Y. Ricky, and J. M. Kerta, “Item Rating and Football Score Priori Prediction Algorithm,” *Int. Conf. Inf. Manag. Technol. ICIMTech*, 2019.

[8] Nur Fadilah and Sigit Priyanta, “Automatic Essay Scoring Using Data Augmentation in Bahasa Indonesia,” *IJCCS Indones. J. Comput. Cybern. Syst.*, vol. 16, no. 4, pp. 401–410, Oct. 2022, doi: https://doi.org/10.22146/ijccs. 76396.

[9] Dongil Kim and Seokho Kang, “Effect of Irrelevant Variables on Faulty Wafer Detection in Semiconductor Manufacturing,” *MDPI Energ.*, vol. 12, Jul. 2019, doi: 10.3390/en12132530.

[10] Achmad Siddik Fathoni and Sri Hartati, “Knowledge-Based Systems Selection of Contraceptive Equipment for The Handling of Uncertainty,” *IJCCS Indones. J. Comput. Cybern. Syst.*, vol. 15, pp. 111–120, Apr. 2021, doi: 10.22146/ijccs.58305.

[11] T. Kim, B. Lieberman, G. Luta, and E. A. Pena, “Prediction Regions for Poisson and Over-Dispersed Poisson Regression Models with Applications in Forecasting the Number of Deaths during the COVID-19 Pandemic,” *Gruyter*, Nov. 2021, doi: https://doi.org/10.1515/stat-2020-0106.

[12] Y. Zhang, W. K. Cheung, and J. Liu, “A Unified Framework for Epidemic Prediction based on Poisson Regression,” *IEEE Trans. Knowl. Data Eng.*, vol. 27, Oct. 2015, doi: Digital Object Identifier no. 10.1109/TKDE.2015.2436918.

[13] T. Inan, “Using poisson model for goal prediction in European football,” *J. Hum. Sport Exerc.*, pp. 1–14, Apr. 2020, doi: https://doi.org/10.14198/jhse.2021.164.16.

[14] Senol Celik, “Predicting the Number of Goals in Football Matches with the Poisson distribution: Example of Spain La Liga,” *Sch. J. Phys. Math. Stat.*, Oct. 2021, doi: 10.36347/sjpms.2021.v08i08.002.

[15] Tobias Fritz, Tomas Gonda, and Paolo Perrone, “De Finetti’s Theorem in Categorical Probability,” Jun. 2021.