International Journal of
Science and Engineering Investigations
vol. 8, issue 88, May 2019
Received on April 27, 2019
ISSN: 2251-8843

Supporter Analysis Using Soccer Momentum Data and Sentiment Feature Quantity

Akira Otsuki[1], Shunta Kogen[2], Masayoshi Kawamura[3]
[1]Nihon University Tokyo Japan
[2]HASEKO COMMUNITY Inc.
[3]MK Future Software
([1]otsuki.akira@nihon-u.ac.jp)

Abstract- This is a transdisciplinary study of sports science and data engineering. Specifically, we used text mining to extract sentiment feature value when supporters were watching the game, and clarified the relation of football (soccer) momentum in the same game as this feature from a quantitative approach. Concretely, It will extraction sentiment feature quantity from SNS data using Word Emotion Polarity Correspondence Table, and will extraction soccer momentum amount data from Football Lab site. Then it will clarify soccer exercise amount data by analysis about these relationships.

Keywords- *Sports Data Analysis, Sentiment Feature Quantity Analysis, Multiple Regression Analysis*

I. INTRODUCTION

In recent years, the use of data in the sports field has been attracting much attention. This is probably due to easier data acquisition and analysis due to technological innovations. By analyzing the acquired data, it is possible to visualize characteristics such as team tendencies, strengths, weaknesses, etc., and to obtain indicators such as tactics, exercise methods, countermeasures against opposing teams, and so on. Under such circumstances, the J League began to release tracking data (momentum data) such as running distance, running speed, etc., from the 2015 season (J League Site: https://www.jleague.jp/match/j1/2016/030501/live/#trackingdata).

Figure 1. Number of SNS users in Japan (ICT Research Institute, 2015)

In addition, due to the recent rapid surge in the use of smartphones, writing on social media such as blogs and Twitter has increased [1-2]. This has made it easier to perform analysis of emotions and sentiments analysis [3] included in such writing using text mining, etc.

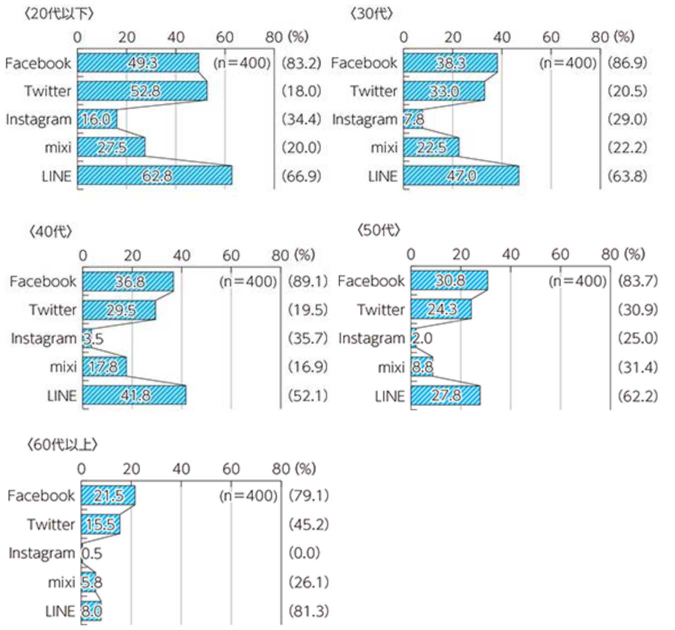


Figure 2. SNS usage rate by age group (ICT Research Institute, 2015)

As shown in Chapter 2, there have been previous studies that analyze tracking data and play data in football or use sentiment analysis, but no study has yet been proposed to analyze the relationship between momentum data and sentiment feature value in football. However, in the Japanese football world, in order to gain new supporters and prevent existing supporters from leaving, it would be very useful to clarify what kind of effect momentum has on supporters, either positive or negative. Therefore, in this study, we designed a model to clarify the relationship between momentum data including tracking data in football, and sentiment feature value.

In the demonstration experiment described in Chapter 4, we obtained momentum data and tweets for 34 games by the Kashima Antlers' in the 2016 Meiji Yasuda J1 League 1st and 2nd Stages, and extracted and analyzed sentiment features. As a result, the momentum that had the most positive influence on Kashima supporters became clear. Further, among the 34 games of Kashima Antlers' first and second stages, the game that had the most positive influence on Kashima supporters was also identified. The results of these analyses are expected to contribute to "supporter management" aimed at gaining new supporters, and preventing existing supporters from leaving, etc.

II. PREVIOUS/RELATED STUDIES

Here, we will summarize previous studies on analysis using football momentum data. In the attack-play optimization algorithm by Guangxiao, et al. [4], an algorithm using Big Data to optimize play leading up to a shoot during an attack, was proposed. Kusano [5] also analyzed trends of shooting position and in-goal position in football. Specifically, he focused on 813 points recorded in the 2010 J1 league, especially 147 points of the top 10 players, By using the series of processes related to these points, that is, points pattern, shooting position, no. of touches up to a shoot, shooting technique and in-goal position, he concluded that to gain points in football, it is important to keep in mind to get into the front area inside the shooter's penalty zone, face the goal, and bring in more precise skills.

Through this analysis, he pointed out that passing which improves the shooting accuracy must be such that the passed ball will enter the attacking side, and when a player is marked, an attack pattern from a side area where it is difficult to obtain the same field of view as the starting point or breakthrough is important by passing a ball from this side area.

Sakaida, et al. [6] et al studied effective tactics for victory by performing game analysis in football, and using six of the best eight teams from Aichi Prefecture in the nationwide high school football preliminaries, analyzed four items,-that is, the ball acquisition zone, the no. of times from acquiring the ball up to a shoot, the no. of passes of attacks leading up to a shoot, and long passes to the front. As a result, they stated that, as regards effective tactics for victory in football, "it is important to get the ball in the attack zone and aim for the goal from fewer pass exchanges", and "actively attack the defending team from the rear with accurate passes".

Recently, there have been more studies using tracking data, and Mizogami, et al. [7] did studies on win/lose factors of Vegalta Sendai using tracking data. Using data such as running distance by position, no. of sprints, no. of shots by time zone and no. of points in all 17 matches played by Vegalta Sendai's 2015 Meiji Yasuda Life J1 League 1st Stage, from statistical analysis and correlation analysis, they found that for Vegarta Sendai to score points and win, it was important to do more sprints than run long distances.

Kurokawa [8] did studies to clarify the influence of no. of sprints and running distance on tactics and win/lose in football, and using tracking data for all 153 matches in Meiji Yasuda Life J1 League 1st and 2nd Stages which can be viewed on "SKY PerfecT!", they compared maximum values of running distance and no. of sprints for both 1st and 2nd stages with the final ranking of each stage, and studied the relationship with the games. As a result, they stated that mere momentum is not enough to lead to victory, rather, victory can be gained by running effectively while thinking about timing and direction.

In a previous study using sentiment analysis, Takahashi, et al. [9], in an analysis of user review information using an "outlier detection method" based on a correlation rule, and using a user review in "Rakuten Travel", applied the outlier detection method of the correlation rule to this, and proposed a method of finding useful reviews.

Honda, et al. [10], in a tweet classification of the burst period in TV programs using sentiment analysis, specified the burst period (the time period when many users post simultaneously), performed tweet classification on sentiments independent of program content, and as a result of evaluations, considered it was possible to classify tweets with high accuracy by dividing opinions about TV into five categories.

Kazumi [11], in the construction of sentiment indices reflecting the stock market and an empirical analysis of stock-price-explanatory power, as a result of performing sentiment analysis only on articles referring to Japanese economic activities, predominantly describe Kazumi's logarithmic return rate and yield, and showed that sentiments then react to the stock market later.

Miyano [12], in a study of the development of facility evaluation tools for creating tourist areas, stated that if a sentiment feature value were extracted from Big Data of tourism and this feature value were applied to create touristic areas, it would be possible to clarify the negative and positive impressions of visitors by area and tourist resources which cannot be judged only by the number of visitors, and it would then be possible to identify the spots which are likely to become more popular in future.

Tokuda, et al. [13], in a sentiment analysis of users who retweet, focused on the change in facial expression that occurs as part of emotional changes at the time of retweeting, and as a result of performing sentiment analysis of users who tweet, showed it was possible to acquire sentiment information by acquiring facial expressions when the button was pressed. In addition, they stated that sentiment analysis could be performed by adapting this method and estimating feelings from the user's facial expression.

In contrast to the previous studies overviewed above, this study collected tweets that were tweeted at the time of a football match, extracts a sentiment feature value, and analyzed the relationship between this feature value and the football momentum of the same game. Consequently, it has novelty that was not found in the previous studies overviewed above.

III. PREPARE YOUR PAPER BEFORE STYLING

A. Analysis data

The subject of this study was all 34 games by the Kashima Antlers in the 2016 Meiji Yasuda J1 League first and second Stages. As data used in the study, we obtained tracking data published in J League (J League Site: https://www.jleague.jp/match/j1/2016/030501/live/#trackingdata.), and football momentum data of the Football LAB (Football Lab site, "2016 Kashima Antlers game schedule and results": http://www.football-lab.jp/kasm/match/?year=2016.)

Next, regarding the method of acquiring tweets, we set the following keywords in the Twitter search API, and got tweets for all the above 34 matches. In addition, if there was something unrelated to football and Kashima Antlers on the tweets, they were deleted manually.

"Kashima OR Antlers OR #antlers date and time*[1]" *[1]Time is set to 2 hours from the start of kickoff

B. Sentiment feature value

We analyzed the tweets obtained in the previous section by Text Mining [14], and extracted bigrams of adjectives and verbs. Sentiment analysis was performed on the extracted bigrams, and the sentiment feature value (SFV) was scored. Specifically, the extracted bigrams were scored against the words sentiment polarity correspondence table shown in Fig.3, and SFV was obtained by the following equation. Fig.4 shows an image of the SFV calculation.

$$\mathrm{SFV} = \frac{\left(\sum_{n=1}^{j} PWS + \sum_{n=1}^{i} NWS\right)}{i+j} \quad (1)$$

where:

I = no. of morphemes in positive words

j = no. of morphemes in negative words

PWS = positive score of morpheme units

NWS = negative score of morpheme units

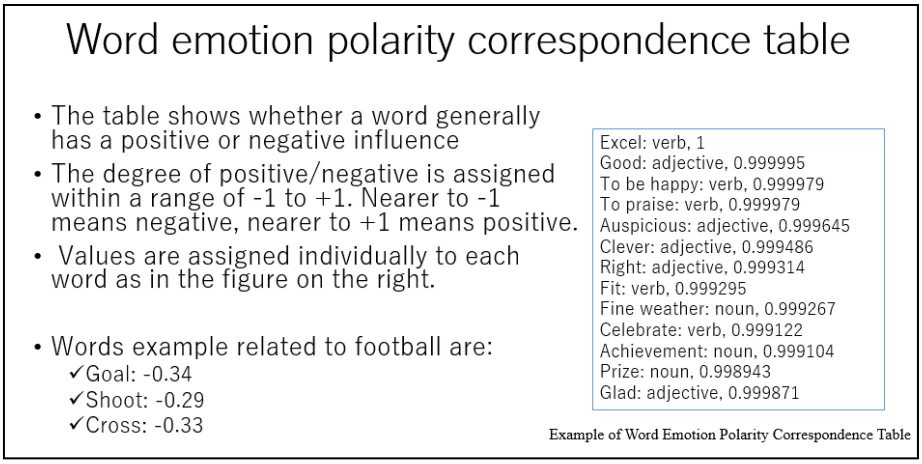


Figure 3. Word emotion polarity correspondence table (created by the author) (Word Emotion Polarity Correspondence Table: http://www.lr.pi.titech.ac.jp/~takamura/pubs/pn_ja.dic.)

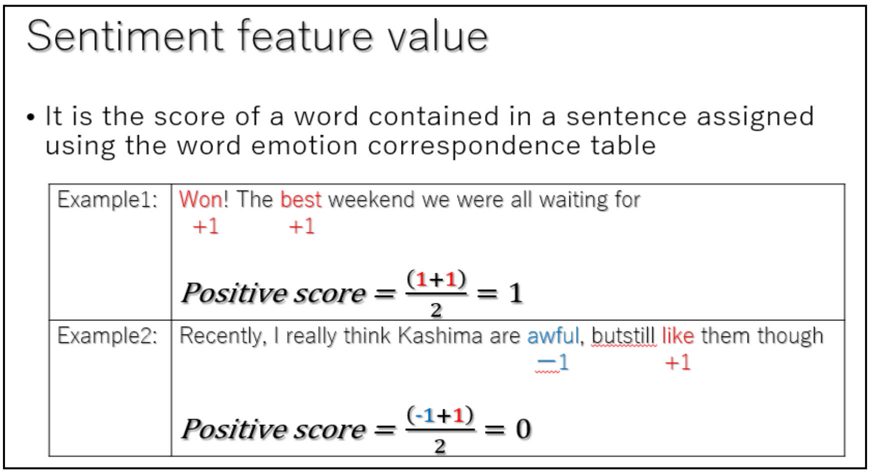


Figure 4. Sentiment feature value (created by author)

IV. SURVEY OF FOOTBALL MOMENTUM CORRELATED WITH SENTIMENT FEATURE VALUE

In this study, in order to analyze the relationship between sentiment feature value and football momentum using multiple regression analysis, explanatory variables set for multiple regression analysis were investigated using correlation analysis.

Specifically, the sentiment feature amount (SFV) in the previous section was set as the objective variable, and as explanatory variables, we decided to set them from those in the "dummy variables" and "team stats (units: times, % and Km)" in Table 1 after examining those which had a correlation with the objective variable.

In this dummy variable, there are two items, i.e., "win/lose" and "home", and as "win/lose", it is 1 if Kashima Antlers won, 0 if it lost, and 0.5 if it drew. In the item "home", 1 was set if it was a home game for Kashima Antlers, and 0 if it was an away game. As the other explanatory variables, tracking data published in the J League and football momentum data of the Football LAB were set as candidates. Specifically, these were as follows:

No. of shots, no. of shots in frame, no. of shots by penalty kicks (PK), no. of passes, no. of crosses, no. of direct free kicks (FK), no. of indirect free kicks (FK), no. of cross kicks (CK), no. of throw-ins, no. of dribbles, no. of tackles, no. of clears, no. of intercepts, no. of offsides, no. of warnings, no. of pitch send-offs, no. of 30 m line intrusions, total running distance, no. of sprints

Then, the correlation coefficient between these explanatory variables and the objective variable was examined. The results are shown in Figs.5 to Fig.24.

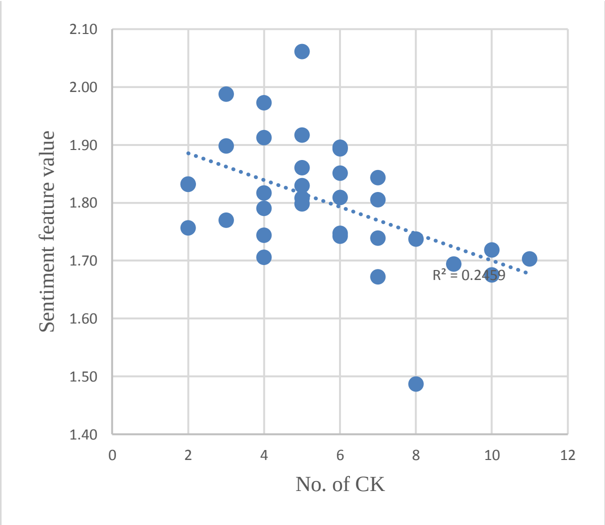


Figure 5. Correlation between sentiment feature value and no. of corner kicks (ck) (correlation coefficient=-0.50)

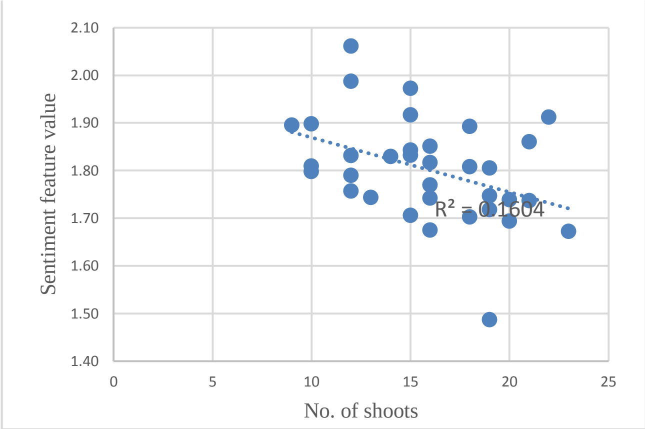


Figure 6. Correlation between sentiment feature value and no. of shoots (*sh*) (correlation coefficient=-0.40)

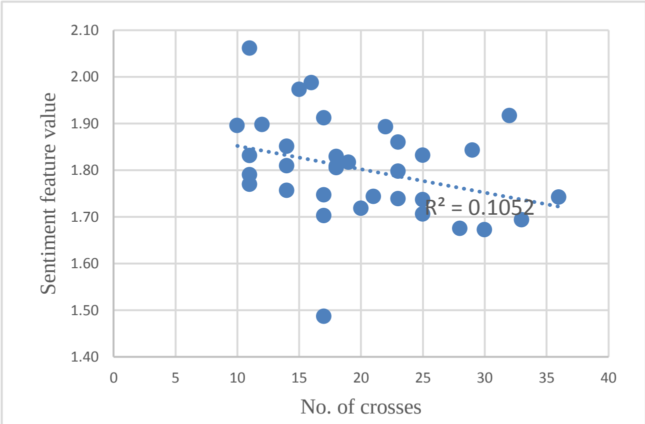


Figure 7. Correlation between sentiment feature value and no. of crosses (*cr*) (correlation coefficient=-0.32）Acknowledgements

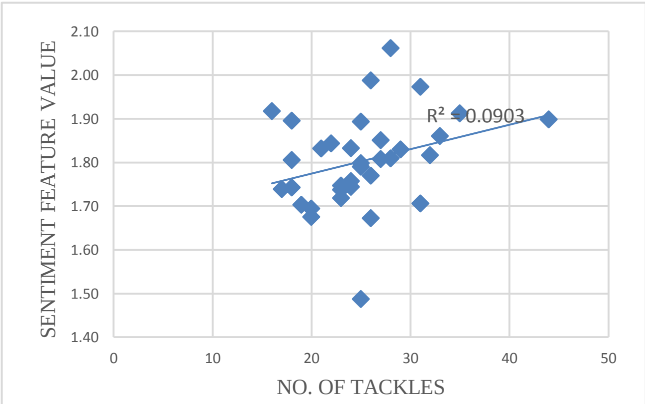


Figure 8. Correlation between sentiment feature value and no. of tackles (*ta*) (correlation coefficient=0.30)

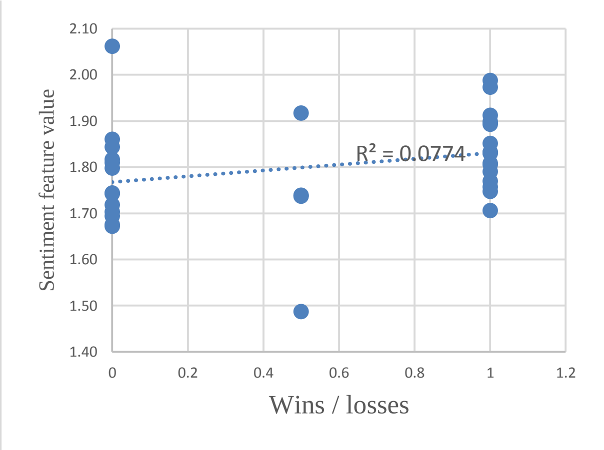


Figure 9. Correlation between sentiment feature value and wins/losses (correlation coefficient=0.28)

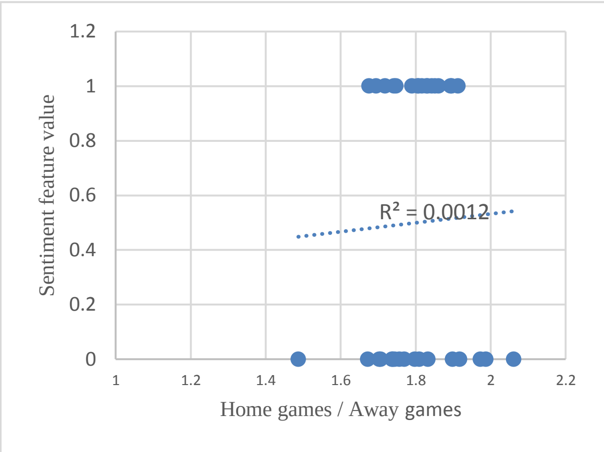


Figure 10. Correlation between sentiment feature value and home games (correlation coefficient=0.03)

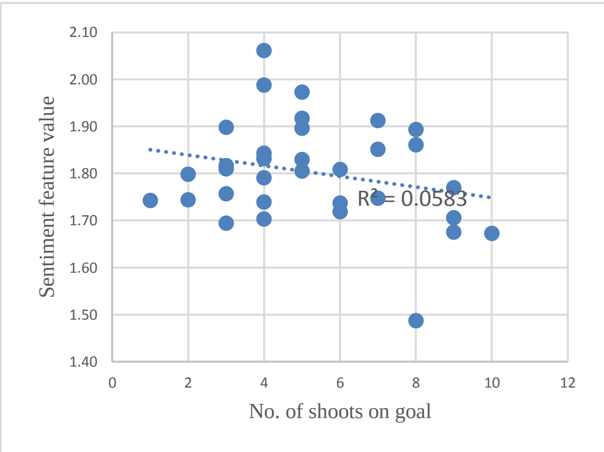


Figure 11. Correlation between sentiment feature value and no. of shoots on goal(correlation coefficient=-0.24)

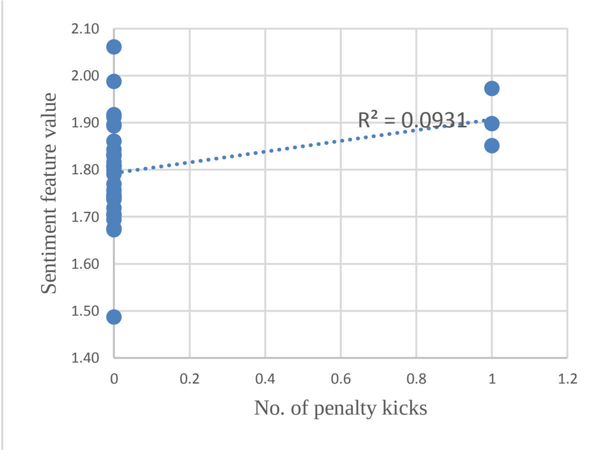


Figure 12. Correlation between sentiment feature value and no. of penalty kicks (correlation coefficient=0.31)

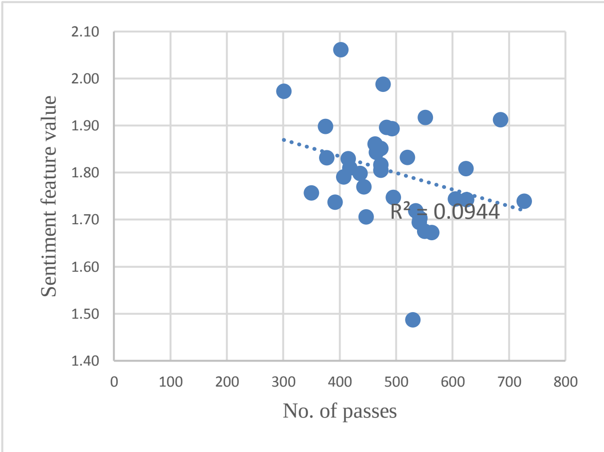


Figure 13. Correlation between sentiment feature value and no. of passes (correlation coefficient=-0.31)

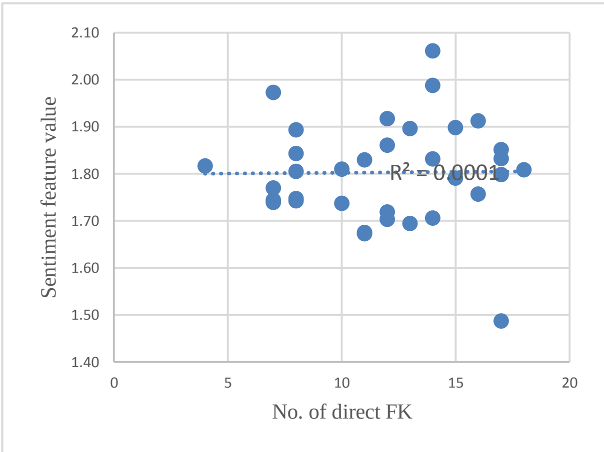


Figure 14. Correlation between sentiment feature value and no. of direct free-kicks (correlation coefficient=0.01)

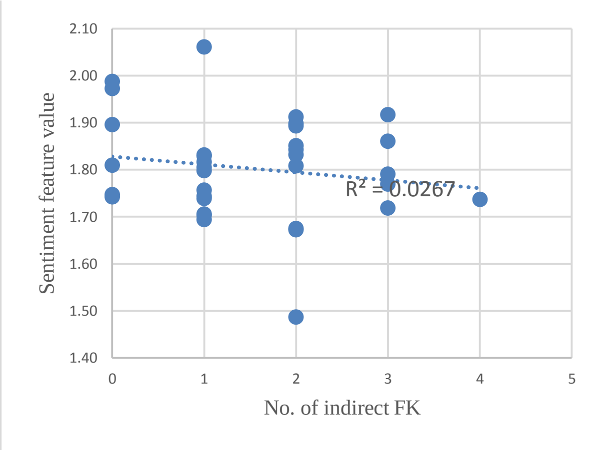


Figure 15. Correlation between sentiment feature value and no. of indirect free-kicks (correlation coefficient=-0.16)

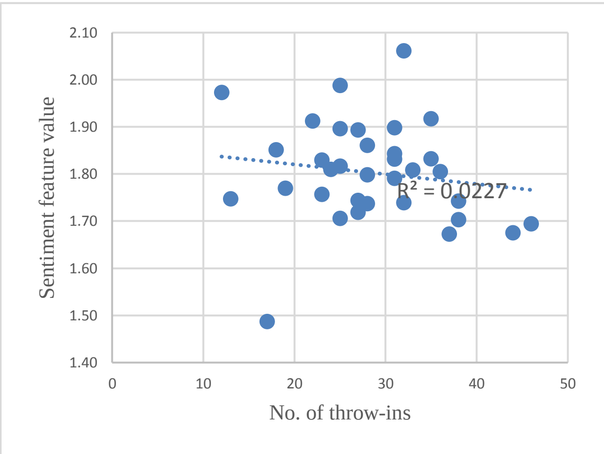


Figure 16. Correlation between sentiment feature value and no. of throw-ins (correlation coefficient=-0.15)

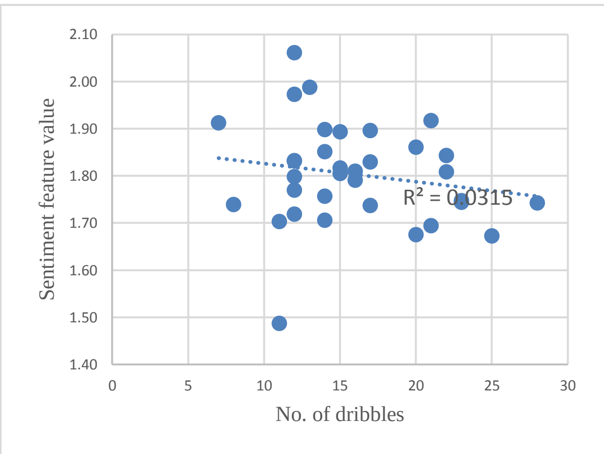


Figure 17. Correlation between sentiment feature value and no. of dribbles (correlation coefficient=-0.18)

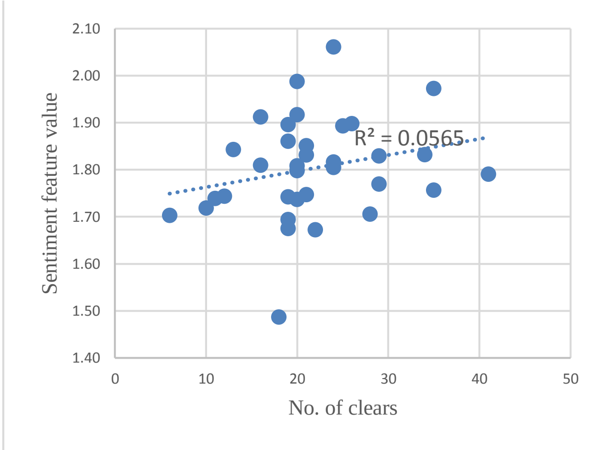


Figure 18. Correlation between sentiment feature value and no. of clears (correlation coefficient=0.24)

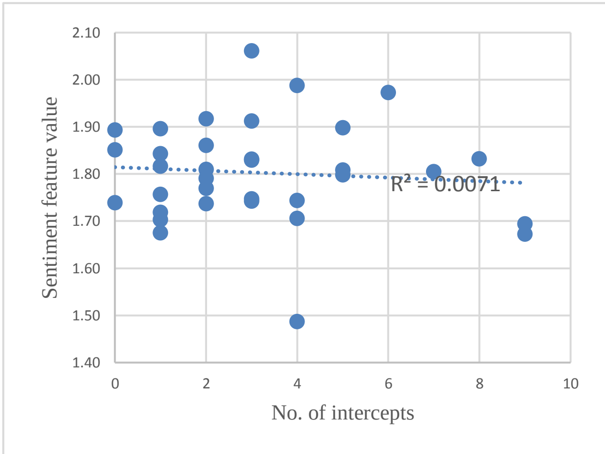


Figure 19. Correlation between sentiment feature value and no. of intercepts (correlation coefficient=-0.08)

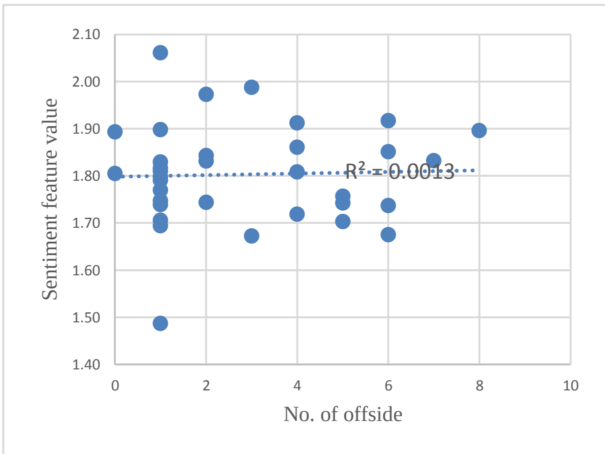


Figure 20. Correlation between sentiment feature value and no. of offside (correlation coefficient=0.04)

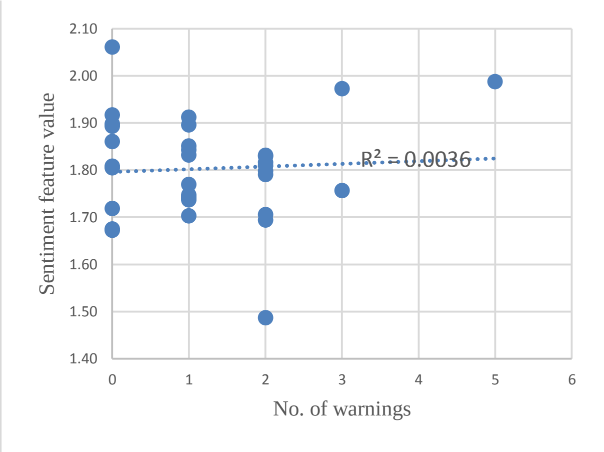


Figure 21. Correlation between sentiment feature value and no. of warnings (correlation coefficient=0.06)

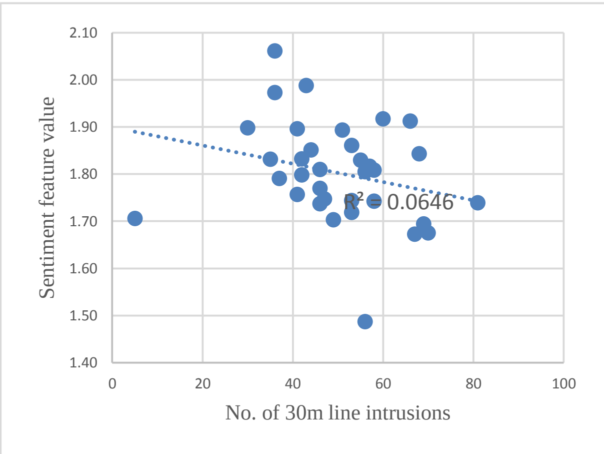


Figure 22. Correlation between sentiment feature value and no. of 30m line intrusions (correlation coefficient=-0.25)

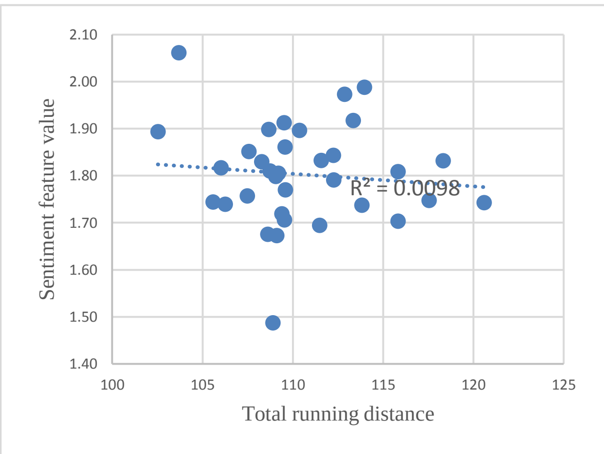


Figure 23. Correlation between sentiment feature value and total running distance (correlation coefficient=-0.10)

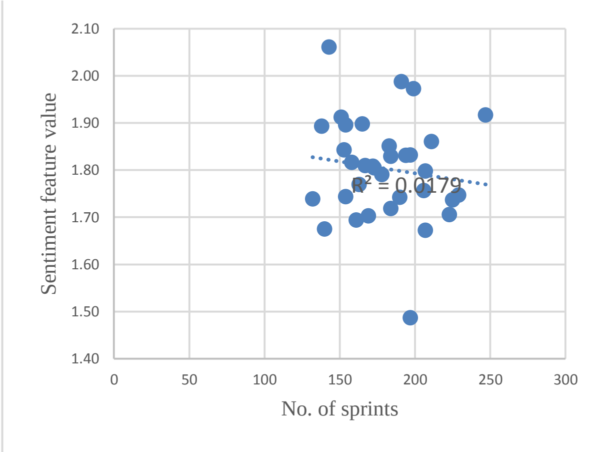


Figure 24. Correlation between sentiment feature value and no. of sprints (correlation coefficient=-0.13)

V. SURVEY OF FOOTBALL MOMENTUM CORRELATED WITH SENTIMENT FEATURE VALUE

According to Peter et al [15], a sample size of about 10 times or more that of the explanatory variables is required. Since the sample size here is all 34 games including the 1st and 2nd Stages of Kashima Antlers in the 2016 Meiji Yasuda J1 League, four of the explanatory variables were considered appropriate.

From the results of the correlation analysis in the previous section, "no. of CK (b_{ck})", "no. of shots (b_{sh})", "no. of crosses (b_{cr})" and "no. of tackles (b_{ta})", that are the momentum data of the TOP 4 for which a correlation with the sentiment feature value was observed, were set as explanatory variables. The multiple regression equation created as described above is shown in Table 1 and Equation 2. Here, b_0 is an explanatory variable, b_{ck} is an explanatory variable corner kick, b_{sh} is an explanatory variable shoot, b_{cr} is an explanatory variable cross, b_{ta} is an explanatory variable tackle, and x_{ck}, x_{sh}, x_{cr}, x_{ta} are partial regression coefficients of each explanatory variable. Finally, e represents the residual error:

$$Y = b_0 + b_{ck}x_{ck} + b_{sh}x_{sh} + b_{cr}x_{cr} + b_{ta}x_{ta} + e \qquad (2)$$

VI. ANALYSIS RESULTS

For data of all 34 games including the 1st and 2nd Stages of Kashima Antlers in the 2016 Meiji Yasuda J1 League, the analysis results obtained by the multiple regression model shown in Equation 1, are shown in Fig.25 and Fig.26.

Fig.25 plots the partial regression coefficients (explanatory variables) of the multiple regression model shown in the previous chapter, and "tackle no." had the highest score. As shown in Fig. 8, since there was a positive correlation with the sentiment feature value, although it was not strong, from the analysis result in this study, it was concluded that tackle no. has a positive influence on supporters.

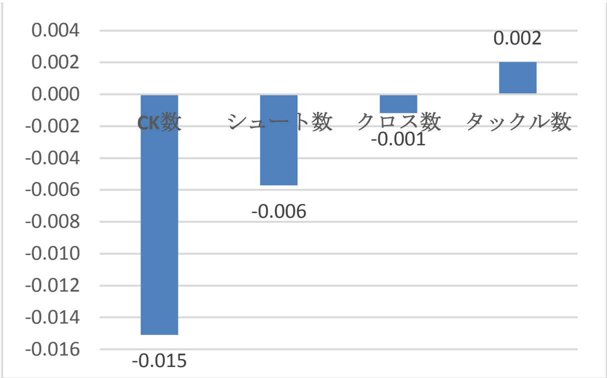
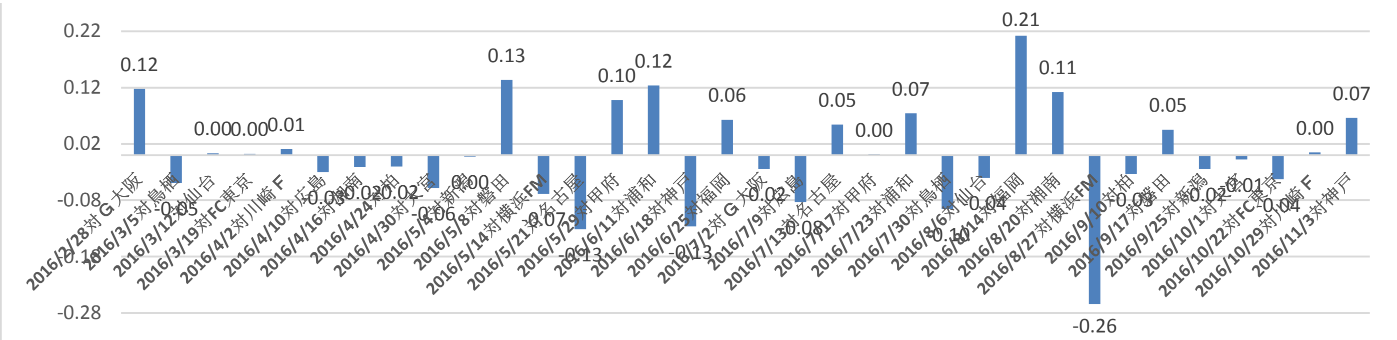


Figure 25. Plot of partial regression coefficients (explanatory variables)

Next, Fig.26 is a graph of the residual error e in the present multiple regression equation"$Y = b_0 + b_{ck}x_{ck} + b_{sh}x_{sh} + b_{cr}x_{cr} + b_{ta}x_{ta} + e$". Since the objective variable Y in this multiple regression equation is the sentiment feature value shown in Section 3.2.1, it can be interpreted that the higher the residual error, the more positive was the influence on supporters than expected.

Fig.26 shows the result of calculating the residual error e mentioned above for all the samples, that is, all 34 games including the 1st and 2nd Stages of Kashima Antlers in the 2016 Meiji Yasuda J1 League. From Fig.26, it was found that the game that had a more positive influence on supporters than expected was the game against Fukuoka held on 14/8/2016, the reason of which is discussed in the next section.



Date	Game (Soccer Club Name)		
28/2/2016	Kashima Antlers	vs	G Osaka
5/3/2016	Kashima Antlers	vs	Tosu
12/3/2016	Kashima Antlers	vs	Sendai
19/3/2016	Kashima Antlers	vs	FC Tokyo
2/4/2016	Kashima Antlers	vs	Kawasaki F
10/4/2016	Kashima Antlers	vs	Hiroshima
16/4/2016	Kashima Antlers	vs	Shonan
24/4/2016	Kashima Antlers	vs	Kashiwa
30/42016	Kashima Antlers	vs	Omiya
4/5/2016	Kashima Antlers	vs	Niigata
8/5/2016	Kashima Antlers	vs	Iwata
14/5/2016	Kashima Antlers	vs	Yokohama FM
21/5/2016	Kashima Antlers	vs	Nagoya
29/5/2016	Kashima Antlers	vs	Kofu
11/6/2016	Kashima Antlers	vs	Urawa
18/6/2016	Kashima Antlers	vs	Kobe
25/6/2016	Kashima Antlers	vs	Fukuoka
2/7/2016	Kashima Antlers	vs	G Osaka
9/7/2016	Kashima Antlers	vs	Hiroshima
13/7/2016	Kashima Antlers	vs	Nagoya
17/7/2016	Kashima Antlers	vs	Kofu
23/7/2016	Kashima Antlers	vs	Urawa
30/7/2016	Kashima Antlers	vs	Tosu
6/8/2016	Kashima Antlers	vs	Sendai
14/8/2016	Kashima Antlers	vs	Fukuoka
20/8/2016	Kashima Antlers	vs	Shonan
27/8/2016	Kashima Antlers	vs	Yokohama FM
10/9/2016	Kashima Antlers	vs	Kashiwa
17/9/2016	Kashima Antlers	vs	Iwata
25/9/2016	Kashima Antlers	vs	Niigata
1/10/2016	Kashima Antlers	vs	Omiya
22/10/2016	Kashima Antlers	vs	FC Tokyo
29/10/2016	Kashima Antlers	vs	Kawasaki F
3/11/2016	Kashima Antlers	vs	Kobe

Figure 26. Residual error graph

TABLE I. MULTIPLE REGRESSION MODEL (THE DUMMY VARIABLE IS DEFINED AS "1" IF KASHIMA ANTLERS WINS, "0" WHEN KASHIMA ANTLERS LOSES, AND "0.5" FOR A DRAW. "1" IS FOR A GAME PLAYED AT HOME, AND 0 FOR A "GAME" PLAYED AWAY).

Date	Opposing team	Objective variable	Dummy variable (Explanatory variable)			Team stats (units: no. of times, % and Km) (Explanatory variables)																		
		Negative/ positive analysis	Win/lose		Home	No. of shots	.o. of shots on goal	No. of PK	No. of passes	No. of crosses	No. of direct FK	No. of indirect FK	No. of CK	No. of throw-ins	No. of dribbles	No. of tackles	No. of clears	No. of intercepts	No. of offside	No. of warnings	No. of send-offs	No. of 30m line intrusions	Total running distance (km)	No. of sprints
28/2/2016	G Osaka	1.99	○1-0	1	0	12	4	0	477	16	14	0	3	25	13	26	20	4	3	5	0	43	113.963	191
5/3/2016	Tosu	1.83	○1-0	1	1	12	4	0	377	11	14	1	2	31	12	21	21	3	2	2	0	35	118.322	194
12/3/2016	Sendai	1.70	●0-1	0	0	18	4	0	542	17	12	1	11	38	11	19	6	1	5	1	0	49	115.829	169
19/3/2016	FC Tokyo	1.81	○2-0	1	1	18	6	0	624	18	18	2	5	33	22	27	20	5	4	0	0	58	115.836	172
2/4/2016	Kawasaki F	1.74	△1-1	0.5	0	21	6	0	392	25	10	4	8	28	17	23	20	2	6	1	0	46	113.82	225
10/4/2016	Hiroshima	1.75	○4-1	1	1	19	7	0	495	17	8	0	6	13	23	23	21	3	1	1	0	47	117.546	229
16/4/2016	Shonan	1.83	○3-0	1	0	15	4	0	520	25	17	2	2	35	12	24	34	8	7	1	0	42	111.572	197
24/4/2016	Kashiwa	1.74	●0-2	0	1	16	1	0	625	36	8	0	6	38	28	18	19	3	5	1	0	58	120.604	190
30/42016	Omiya	1.67	●0-0	0	0	23	10	0	563	30	11	2	7	37	25	26	22	9	3	0	0	67	109.115	207
4/5/2016	Niigata	1.83	○2-1	1	1	14	5	0	415	18	11	1	5	23	17	29	29	3	1	2	0	55	108.284	184
8/5/20168	Iwata	1.92	△1-1	0.5	0	15	5	0	552	32	12	3	5	35	21	16	20	2	6	0	0	60	113.346	247
14/5/2016	Yokohama FM	1.79	○1-0	1	1	12	4	0	407	11	15	3	4	31	16	25	41	2	1	2	0	37	112.266	178
21/5/2016	Nagoya	1.71	○3-2	1	0	15	9	0	447	25	14	1	4	25	14	31	28	4	1	2	0	5	109.54	223
29/5/2016	Kofu	1.91	○4-0	1	1	22	7	0	685	17	16	2	4	22	7	35	16	3	4	1	0	66	109.523	151
11/6/2016	Urawa	1.97	○2-0	1	0	15	5	1	301	15	7	0	4	12	12	31	35	6	2	3	0	36	112.873	199
18/6/2016	Kobe	1.76	○2-1	1	0	12	3	0	350	14	16	1	2	23	14	24	35	1	5	3	0	41	107.477	206
25/6/2016	Fukuoka	1.90	○2-0	1	1	9	5	0	483	10	13	0	6	25	17	18	19	1	8	1	0	41	110.378	154
2/7/2016	G Osaka	1.82	●1-3	0	1	16	3	0	473	19	4	1	4	25	15	32	24	1	1	2	0	57	106.021	158
9/7/2016	Hiroshima	1.77	○4-2	1	0	16	9	0	443	11	7	3	3	19	12	26	29	2	1	1	0	46	109.582	163
13/7/2016	Nagoya	1.80	○3-0	1	1	19	5	0	473	18	8	1	7	36	15	18	24	7	0	0	0	56	109.203	173
17/7/2016	Kofu	1.74	△3-3	0.5	0	20	4	0	727	23	7	1	7	32	8	17	11	0	1	1	0	81	106.255	132
23/7/2016	Urawa	1.84	●1-2	0	1	15	4	0	465	29	8	2	7	31	22	22	13	1	2	1	0	68	112.253	153
30/7/2016	Tosu	1.74	●0-1	0	0	13	2	0	605	21	7	1	4	27	23	24	12	4	2	1	0	53	105.577	154
6/8/2016	Sendai	1.68	●0-1	0	1	16	9	0	551	28	11	2	10	44	20	20	19	1	6	0	0	70	108.612	140
14/8/2016	Fukuoka	2.06	○□2-1	0	0	12	4	0	402	11	14	1	5	32	12	28	24	3	1	0	0	36	103.685	143
20/8/2016	Shonan	1.89	○1-0	1	1	18	8	0	493	22	8	2	6	27	15	25	25	0	0	0	0	51	102.53	138
27/8/2016	Yokohama FM	1.49	△2-2	0.5	0	19	8	0	530	17	17	2	8	17	11	25	18	4	1	2	0	56	108.907	197
10/9/2016	Kashiwa	1.81	●0-2	0	0	10	3	0	418	14	10	0	6	24	16	28	16	2	1	2	0	46	108.738	167
17/9/2016	Iwata	1.85	○3-0	1	1	16	7	1	473	14	17	2	6	18	14	27	21	0	6	1	0	44	107.567	183
25/9/2016	Niigata	1.90	○2-0	1	0	10	3	1	375	12	15	2	3	31	14	44	26	5	1	0	0	30	108.666	165
1/10/2016	Omiya	1.69	●1-3	0	1	20	3	0	541	33	13	1	9	46	21	20	19	9	1	2	0	69	111.487	161
22/10/2016	FCTokyo	1.80	●1-2	0	0	10	2	0	436	23	17	1	5	28	12	25	20	5	1	2	0	42	109.059	207
29/10/2016	Kawasaki F	1.72	●0-1	0	1	19	6	0	535	20	12	3	10	27	12	23	10	1	4	0	0	53	109.394	184
3/11/2016	Kobe	1.86	●0-1	0	1	21	8	0	463	23	12	3	5	28	20	33	19	2	4	0	0	53	109.574	211

VII. DISCUSSION OF ANALYSIS RESULTS

Based on the analysis results in the previous section, it was revealed that the game whose SFV was higher than expected, was the match against Fukuoka which was held on 14/8/2016. This can be explained as follows. As shown in Table2, the team lost 3 consecutive League games until the match against Fukuoka on 14/8/2016, the victory against Fukuoka after consecutive defeat in the above-mentioned game had a major positive influence on the positive score. This can also be appreciated by looking at the transition of SFV in Table3.

TABLE II. SFV AND WINS/DEFEATS FOR LAST 3 GAMES PRECEDING MATCH AGAINST FUKUOKA ON 14/8/2016

Date	Game (Soccer Club Name)	SFV	Win / Lose
23/7/2016	Kashima Antlers vs Urawa	1.84	●1-2
30/7/2016	Kashima Antlers vs Tosu	1.74	●0-1
6/8/2016	Kashima Antlers vs Sendai	1.68	●0-1
14/8/2016	Kashima Antlers vs Fukuoka	2.06	○2-1

Table3 shows the transition of SFV when they won after two consecutive losses in the same season, but at this time (after two consecutive losses), it is seen that SFV was also increasing. Due to this, it appears that the residual error value of SFV in the match against Fukuoka was higher than anticipated since the team was able to stop a run of losing streaks, and achieved victory after such a long time. When we examined several related tweets manually, we were able to confirm tweets asserting the reasons mentioned above.

TABLE III. TREND OF SFV WHEN WINNING AFTER A SERIES OF LOSSES IN THE SSAME SEASON

Date	Game (Soccer Club Name)	SFV	Win / Lose
24/4/2016	Kashima Antlers vs Kashiwa	1.74	●0-2
30/4/2016	Kashima Antlers vs Omiya	1.67	●0-0
4/5/2016	Kashima Antlers vs Niigata	1.83	○2-1

From the analysis results in this study, it was clear that not only the victorious outcome, but also the increase in the number of tackles had a positive influence on Kashima Antlers supporters. In addition, considering the result of the match against Fukuoka, it can be seen that supporters are influenced by victory or defeat, so a win by Kashima Antlers is still what supporters desire most, and if this continued, it would surely lead to getting new supporters.

In addition, it was found that Kashima supporters often have negative emotions against crosses and corner kicks. As a result of manually examining tweets that contained "cross" or "corner kick", however, most of them had opinion that "cross" or "corner kick" itself was not often linked to scoring opportunities or actual scores and could not be expected to lead to a win. Therefore, by improving the precision of cross or corner kick, and improving it so it has a positive influence on supporters, it would be possible to induce greater excitement than before, prevent supporters from leaving, and lead to an increase in the number of spectators.

VIII. CONCLUSION

The recent use of data in the sports field has been attracting much attention, and the J League is now releasing tracking data from the 2015 season. In addition, due to the recent rapid surge in the penetration rate of smartphones, blogging on social media has increased, and it is now easy to perform sentiment analysis to analyze emotions expressed in writing.

Studies analyzing play data in football and papers using sentiment analysis have appeared, but so far, there have been no studies using the two. However, in the world of Japanese football it is considered extremely important to do this research to gain new supporters, and prevent supporters from leaving.

Therefore, in this study, we collected tweets posted on Twitter during Kashima Antlers matches for all 34 games including the Meiji Yasuda J1 League in FY2016. We then analyzed the acquired data by morphological analysis to extract bigrams of adjectives and verbs, and performed sentiment analysis to score sentiment features. Next, correlation analysis was performed with sentiment feature value and football momentum value data, the four football momentum amounts of "cross", "shoot", "corner kick" and "tackle" which had a correlation with the sentiment feature value were set as explanatory variables, and multiple regression analysis was performed by setting the sentiment feature value as the objective variable. As a result of this analysis, it became clear that the momentum amount that had the most positive influence on Kashima Antlers supporters was "tackle". In addition, it was found that the game that had the most positive influence on supporters over expectations was the game against Fukuoka held on 14/8/2016.

As a future topic of this research, improvement of the word emotion polarity correspondence table can be considered. The cells shown in blue in Fig. 29 are games where the SFV is low although the team was victorious, and the cells shown in yellow are games with high SFV although the team lost, so it is conceivable that the word emotion polarity correspondence table was for general use, and not specially made for football. Therefore, in the future, we would like to arrange the word emotion polarity correspondence table for football, and increase the precision of this analysis method.

TABLE IV. SUBJECT OF THIS STUDY

Date	Game (Soccer Club Name)	SFV	Win / × Lose
28/2/2016	Kashima Antlers vs G Osaka	1.99	○1-0
5/3/2016	Kashima Antlers vs Tosu	1.83	○1-0
12/3/2016	Kashima Antlers vs Sendai	1.70	●0-1
19/3/2016	Kashima Antlers vs FC Tokyo	1.81	○2-0
2/4/2016	Kashima Antlers vs Kawasaki F	1.74	△1-1
10/4/2016	Kashima Antlers vs Hiroshima	1.75	○4-1
16/4/2016	Kashima Antlers vs Shonan	1.83	○3-0
24/4/2016	Kashima Antlers vs Kashiwa	1.74	●0-2
30/42016	Kashima Antlers vs Omiya	1.67	●0-0
4/5/2016	Kashima Antlers vs Niigata	1.83	○2-1
8/5/20168	Kashima Antlers vs Iwata	1.92	△1-1
14/5/2016	Kashima Antlers vs Yokohama FM	1.79	○1-0
21/5/2016	Kashima Antlers vs Nagoya	1.71	○3-2
29/5/2016	Kashima Antlers vs Kofu	1.91	○4-0
11/6/2016	Kashima Antlers vs Urawa	1.97	○2-0
18/6/2016	Kashima Antlers vs Kobe	1.76	○2-1
25/6/2016	Kashima Antlers vs Fukuoka	1.90	○2-0
2/7/2016	Kashima Antlers vs G Osaka	1.82	●1-3
9/7/2016	Kashima Antlers vs Hiroshima	1.77	○4-2
13/7/2016	Kashima Antlers vs Nagoya	1.80	○3-0
17/7/2016	Kashima Antlers vs Kofu	1.74	△3-3
23/7/2016	Kashima Antlers vs Urawa	1.84	●1-2
30/7/2016	Kashima Antlers vs Tosu	1.74	●0-1
6/8/2016	Kashima Antlers vs Sendai	1.68	●0-1
14/8/2016	Kashima Antlers vs Fukuoka	2.06	○2-1
20/8/2016	Kashima Antlers vs Shonan	1.89	○1-0
27/8/2016	Kashima Antlers vs Yokohama FM	1.49	△2-2
10/9/2016	Kashima Antlers vs Kashiwa	1.81	●0-2
17/9/2016	Kashima Antlers vs Iwata	1.85	○3-0
25/9/2016	Kashima Antlers vs Niigata	1.90	○2-0
1/10/2016	Kashima Antlers vs Omiya	1.69	●1-3
22/10/2016	Kashima Antlers vs FC Tokyo	1.80	●1-2
29/10/2016	Kashima Antlers vs Kawasaki F	1.72	●0-1
3/11/2016	Kashima Antlers vs Kobe	1.86	●0-1

REFERENCES

[1] ICT Research Institute, "Survey of SNS Usage Trends in Fiscal 2015": http://ictr.co.jp/report/20150729000088-2.html.

[2] Ministry of Internal Affairs and Communications, Heisei 27th Edition Information Communication White Paper, "Using SNS", http://www.soumu.go.jp/johotsusintokei/whitepaper/en/h27/html/nc242220.html.

[3] Hashimoto Kazuyuki et al, "Reputation trend extraction from microblogging by sentiment analysis and topic extraction", Transactions of the Institute of Electronics, Information and Communication Engineers D, Information and Systems 94 (11), pp. 1762-1772, 2011.

[4] Guangxiao Xu, Keiryo Osawa, Shingotomi Shota, Teruo Ando, Hiroya Suzuki, Naohiko Nishijima, "Development of play optimization algorithm in football attacks, "Special attention" New challenges in sports statistical science", [Research notes] Statistical Mathematics Vol. 65, No. 2, pp. 309 - 321, 2017.

[5] Kusui Kusano, "Analysis of trends in shot points and goal-in points in football - at "J-League Division 1 in 2010" in the top 10 J Leaguers -", Bulletin of Sendai University, Vol. 44, No. 1: 31-41, 2012.

[6] Masataka Sakaida, Hiroyuki Taki, Nobutaka Kito, "On effective tactics for victory as seen from game analysis in football - from High School Football Championship in Aichi Prefecture Preliminaries", Bulletin of Aichi Gakuin University Education Department 54, 2 Issue, pp. 49-59, 2006.

[7] Takeshi Mizogami, Hayato Waizumi, Kyoko Shirahata, Hideyuki Yoshii, "A study of the factors in winning and losing in Vegalta Sendai using tracking data", Bulletin of Sendai University, 47 (2), pp 49-56, 2016.

[8] Hiroku Kurokawa, "The influence of the number of sprints and total running distance on tactics and win points in football", Abstracts of Biwako Seikei Sports College Graduate Research, 2016.

[9] Tsuyoshi Takahashi, Toshiyuki Tenkasu, Hiroyuki Kitagawa, "Analysis of user review information using outlier detection method based on correlation rule", Proceedings of Information Processing Society Conference, Vol. 73, No. 1, pp. 1589-1590,2011.

[10] Haruki Honda, Yokoi Ken, "Tweet classification of burst period in television programs using sentiment analysis", Proceedings of the 76th National Convention of Information Processing Society, 2014.

[11] Takuro Kazumi, "An empirical analysis of the establishment of a sentiment index reflecting the stock market and stock-price-explanatory power", Osaka University Economics, 66 (3), pp. 24 - 36,2016.

[12] Yukitake Miyano, "A Study of the Development of Measuring for Evaluating Tourist Areas: Through Empirical Analysis by Large Tourism Data", Bulletin of Oita Prefectural Art and Culture Junior College, 54, pp.167-180, 2017.

[13] Takuya Tokuda, Ken Yokoi, "Sentiment analysis of retweeting focusing on facial expressions of retweet users", The 76th Annual Conference Papers Collection, pp. 147-148,2014.

[14] Steven Bird, Ewan Klein, Edward Loper, "Introductory Book of Natural Language Processing", O'Reilly, 2010.

[15] Peter Peduzzi et al, "Importance of events per independent variable in proportional hazards regression analysis II. Accuracy and precision of regression estimates", Journal of Clinical Epidemiology, Volume 48, Issue 12, pp 150-151, 1995.

Akira Otsuki Received his Ph.D. in engineering from Keio University (Japan) in 2012. He is currently professor at Nihon University College of Economics (Japan) and Officer at Japan society of Information and knowledge (JSIK). His research interests include Machine learning, Data Mining, Text Mining and Big Data Engineering. Received his Best paper award 2012 at JSIK, received his award in Editage Inspired Researcher Grant in 2012 and received his award in 2018 Kanto branch Award at The Society of Socio-Informatics.

Masayoshi Kawamura is a system engineer (Japan). He received M.S. degree from Kyoto Institute of Technology (Japan) in 1998. His research interests include image processing, digital signal processing, and statistical data analysis.

How to Cite this Article:

Otsuki, A. & Kawamura, M. (2019) Supporter Analysis Using Soccer Momentum Data and Sentiment Feature Quantity. International Journal of Science and Engineering Investigations (IJSEI), 8(88), 1-11. http://www.ijsei.com/papers/ijsei-88819-01.pdf