Classification of texts describing baseball batting results on Twitter using BERT

Kazuma Sakamoto (Department of Production Systems Engineering and Sciences, Komatsu University, kazuma.sakamoto@komatsu-u.ac.jp, Japan)
Riku Ikeda (Graduate School of Sustainable Systems Science, Komatsu University, 22211002@komatsu-u.ac.jp, Japan)
Yoshihiro Ueda (Department of Production Systems Engineering and Sciences, Komatsu University, yoshihiro.ueda@komatsu-u.ac.jp, Japan)

Abstract

With the improvement of communication speeds and the spread of smartphones and IoT devices, there is an abundance of sports video content on the Internet, and users can watch game broadcasts and highlight videos without any time or location restrictions. Although sports games are very long and difficult to watch continuously, users find it more attractive and valuable to watch them in real-time than to record them. In this research, we develop a breaking news system that automatically identifies a specific player's cue, results, and chance scenes by understanding the context of real-time tweets on Twitter and sports commentary from commentators using BERT. This system will enable users who have difficulty in watching in real-time or who are only interested in specific players to watch, as well as to automate breaking news that is currently done manually.

Key words

BERT, Twitter, text classification, sport, baseball

1. Introduction

With the improvement of communication speeds and the spread of smartphones and IoT (Internet of Things) devices, content streaming services known as OTT (Over The Top) (Canadian Radio-television and Telecommunications Commission, 2011), which enable access to programs over the Internet without the need to use dedicated streaming facilities or networks such as cable or satellite are becoming increasingly popular. This allows users to watch live game broadcasts and highlights without any time or location restrictions. In December 2022, Ministry of Economy, Trade and Industry and Japan Sports Agency (2022) reported the "Sports DX Report," which states that sports games are one of the few video contents that are worth watching in real-time because of the live and one-of-a-kind nature of the sport. In addition, the sense of unity among fans is also highly valued, and users find it more attractive and valuable to watch them in real-time than to record them. However, it has its issues. One of them is the length of the games. For example, in the NBA (National Basketball Association), the average game length in the 2021/2022 season was 2 hours and 13 minutes (The Hoops Geek, 2022), and in the MLB (Major League Baseball), it was 3 hours and 3 minutes in the 2022 season (Statista, 2022). In addition to the length of the games, there are users who cannot watch games in real-time due to work, school, or time differences. MLB, in particular, is the only non-timed sport among the Big Four (NFL, MLB, NBA, and NHL). In order to shorten game length, MLB has implemented a "Pitch Timer" (also known as "Pitch Clock") in 2023 (MLB Advanced Media, 2023). It is hoped that this will prevent fans from losing interest. However, since it is a non-timed game, there is a limit to shorten game length.

One of the methods to solve these issues is the need to keep track of the status of sports games in real-time and encourage users to watch sports games efficiently. Twitter is being used as a means to keep track of the status of sports games. Twitter is what's happening and what people are talking about right now (X Corp., 2023). Twitter has excellent real-time performance and ability to spread information. Previous research on Twitter includes those for observing the real world by considering Twitter users as social sensors (Sasaki and Matsuo, 2012; Sakamoto et al., 2018), detecting earthquakes (Sasaki, 2010; Toriumi and Baba, 2016), and using it for road management (Fujimoto et al, 2018). Twitter plays a very important role in disasters such as earthquakes and accidents as a means to understand the situation before rescue teams and the press arrive at the scene. These studies used SVM (Support Vector Machine), keyword, and retweet-based clustering methods. There have also been studied to estimate the status from tweets containing the status of sports games (Fujimoto and Ushiama, 2021). However, this research relies on the number of tweets, and SVM is mainly used for two-category classification, and there is a problem that the classification results change significantly as the number of categories increases.

In this research, we focus on the field of sports, where there is a wealth of information on breaking news, and we aim to develop a system for MLB that notifies users of a specific player's cue, results, and chance scenes. In order to develop such a system, we first focus on extracting information in real time. Similar to previous research, this research also considers Twitter as a social sensor and proposes a method to classify each scene from a variety of tweets to understand the status of sports games. This will contribute to the development of breaking news system.

The structure of this paper is as follows: Section 2 provides an overview of the system developed in this research. Section

3 describes the evaluation of the system using tweets and classification accuracy, Section 4 examines the versatility of this system, and Section 5 discusses the conclusions.

In this research, tweets are handled in accordance with Twitter Privacy Policy (X Corp., 2023).

2. Method

This section discusses a method for classifying tweets containing information about the status of sports games into set categories and calculating the classification accuracy of the tweets. Figure 1 shows the system flow of this research. As shown in the system flow, the system consists of five functions: "Pre-training," "Tweet Pre-processing," "Annotation," "Fine-tuning," and "Tweet Classification."

2.1 Pre-training

In this function, we use Japanese Wikipedia as text data on the Web and construct a model that has learned Japanese sentence structure by pre-training BERT (Bidirectional Encoder Representations from Transformers) (Devlin et al., 2018). In this research, we use a pre-trained BERT model on Japanese Wikipedia released by the Inui Lab at Tohoku University (Tohoku NLP Group, 2019).

2.1.1 BERT

BERT is a language representation model designed to pre-train deep bidirectional representations using the Transformer (Vaswani et al., 2017). One of the features of BERT is its versatility. Since unlabeled text data is used for pre-training, it is relatively easy to collect a large amount of data. Pre-trained BERT models can be fine-tuned by adding a single output layer, making it possible to create state-of-the-art models for a wide range of downstream NLP (Natural Language processing) tasks.

2.2 Tweet preprocessing

This function cleans noisy tweets that contain noise. By removing the noise specific to tweets, better features and meanings can be extracted, which greatly affects the accuracy of the subsequent process, fine-tuning. First, in this research, removing URLs, question marks, exclamation points, and at signs, and then splitting tweet into tokens using a tokenizer. In addition, separating tweet into words using a MeCab, and then splitting words into tokens using a WordPiece. An example of tokenization is shown in Figure 2.

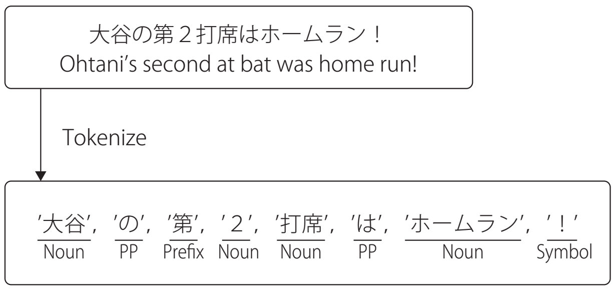

Figure 2: Example of tokenization

Note: PP = Postpositional Particle.

2.3 Annotation

This function creates a dataset from tokenized tweets. First, it extracts tweets containing the statuses of sports game using specific keywords. Next, each extracted tweet is assigned a label for fine-tuning. The labels are assigned a number from 0 to the given category. Fine-tuning is explained in detail in

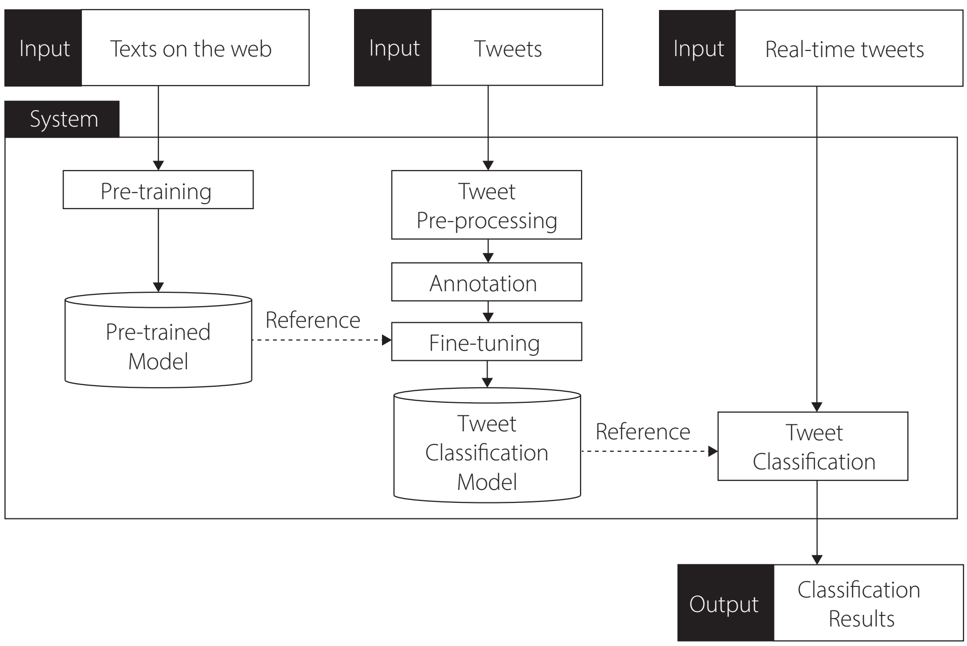

Figure 1: System flow

Section 2.4, and the dataset is explained in detail in Section 3.1.

2.4 Fine-tuning

This function does fine-turning to make BERT specialized for the tweet classification task. Fine-tuning is a method of retraining a model by inputting additional data to a pre-trained model. This allows the model to be specialized for a specific task. It constructs a tweet classification model by referring to a pre-trained model and fine-tuning BERT to specialize in the tweet classification task, using labeled data created by the annotation function as input.

2.5 Tweet classification

This function constructs a BERT model specifically for the tweet classification task. First, fine-tuning is performed by inputting labeled data generated by the annotation function into the pre-trained BERT model. At this point, the labeled data is divided into training data, validation data, and test data. The training and validation data are used to fine-tune BERT, and the test data is used to evaluate BERT performance.

3. Evaluation

3.1 Dataset for fine-tuning

For the dataset used in this research, we selected 59 Twitter accounts that tweeted a relatively large number of baseball game statuses in Japanese and collected their tweets using Twitter API. Therefore, no scraping was performed. The number of collected tweets is 158,667. In this research, we use five categories: "Flyout," "Base on balls," "Groundout," "Home run," and "Strikeout." These categories are explained in detail in Section 3.3. We annotate for these categories. However, when the annotations were performed, extraction using only these keywords resulted in the extraction of some incorrect tweets. These were excluded by visual inspection or by searching for common keywords. Table 1 shows some examples of tweets with only the URLs removed.

Fine-tuning BERT using the completed labeled dataset. Here, in order to take advantage of the characteristic of BERT that allows fine-tuning with a small amount of labeled data, the dataset is created with a small number of tweets. Each labeled dataset is populated with 100 tweets each, for a total of 500 tweets. Table 2 shown that dataset in this research.

3.2 Implementation environment

The implementation environment used in this research is shown in Table 3. Since this research is implemented on a local machine, it is necessary to install PyTorch corresponding for the CUDA version. Other libraries were installed with reference to Introduction to Natural Language Processing with BERT: Practical Programming Using Transformers (Ohmi et al., 2021) and Stockmark's repositories on GitHub (Stockmark, 2022).

3.3 Category

In this research, the five categories introduced in Section 3.1 were set as the classification categories. The reason why these categories were chosen is because of the characteristics of baseball. According to the OFFICIAL BASEBALL RULES 2023 Edition (The Office of the Commissioner of Baseball, 2023), the offensive team's objective is to have its batter become a runner, and its runners advanced. On the other hand, the defensive team's objective is to prevent the offensive team's objective. In other words, batting result is whether the batter becomes a runner or an out. There are thirteen cases listed for a batter to become a runner, which can be summarized as follows: "Hit," "Home run," "Base on balls," "Hit by pitch," "Defensive error," and "Catcher interference." On the other hand, there are fifteen cases listed for a batter to be out, which can be summarized as follows: "Flyout," "Strikeout," "Groundout," and "Batter or runner interference." From these, five main cases were chosen as categories as batting results. See also the OFFICIAL BASEBALL RULES 2023 Edition for more information on the five categories.

3.4 Input tweets

The tweets were randomly selected from the remaining annotated tweets after excluding the tweets used for fine-tuning. Specifically, we selected 100 tweets from each category for input into the system.

3.5 Results

The Precision, Recall, and F-measure for tweets classified by BERT, which were fine-tuned to be specific to the tweet classification task, are shown in Table 4. As shown in Table 4, the F-measure exceeded 0.8 points in the four categories of "Flyout," "Base on balls," "Groundout," and "Strikeout." These factors can be attributed firstly to the fact that we visually checked the tweets when annotating the data for the fine-tuning, which allowed us to create good quality labeled data.

The F-measure for "Home run" was slightly lower than the other categories. These factors can be attributed to the fact that there were only five categories at this stage, and because the "Others" category had not been set up, almost all ambiguous tweets were categorized as "Home run". In the labeled data used for fine-tuning, tweets about home runs had a wide variety of expressions and many of the words used in other categories may have been used. As a measure for improvement, first create an "Others" category. Then, we believe that the number of fine-tuning data can be increased. Examples of misclassified tweets are shown in Table 5.

3.6 Relation to other large language models

BERT is one of the LLMs (Large Language Models); other

Table 1: Tweets example

Keyword	Language	Tweets example
Flyout	Japanese	#大谷翔平 第6打席はレフトへの犠牲フライ！ @ABEMA で視聴中
	Literal Translation	#ShoheiOhtani 6th at bat was a sacrifice fly to left! Watching on @ABEMA
	Japanese	［速報］大谷選手の第3打席ライトフライ @ABEMA で視聴中
	Literal Translation	[Breaking news] Ohtani's third at bat was a flyout to light watching on @ABEMA
	Japanese	コーラ、ハンバーガー、フライドポテト←なにか1つ追加して最強パーティ作れ
	Literal Translation	Coke, hamburger, french fries - add one of these to make the best party.
Base on balls	Japanese	大谷翔平選手の第2打席はフォアボール！①内野安打②四球＃大谷翔平 #MLB ＃野球 ＃エンゼルス
	Literal Translation	Shohei Ohtani's second at bat was a base on balls! 1) Infield hit 2) base on balls #Shohei Ohtani #MLB #Baseball #Angels
	Japanese	［速報］大谷選手の第1打席四球 @ABEMA で視聴中
	Literal Translation	[Breaking news] Ohtani's first at bat was a base on balls watching on @ABEMA
	Japanese	大谷翔平選手！ 3回まで3安打5奪三振無四球0失点球数43 ＃大谷翔平
	Literal Translation	Shohei Ohtani! 3 innings, 3 hits, 5 strikeouts, no base on balls, 0 runs, 43 pitches #ShoheiOhtani
Groundout	Japanese	大谷翔平選手の第1打席はショートゴロ！①遊ゴロ＃大谷翔平 #MLB ＃野球 ＃エンゼルス
	Literal Translation	Shohei Ohtani's first at bat was a groundout to shortstop! 1) Yuugoro #ShoheiOhtani #MLB #Baseball #Angels
	Japanese	［速報］大谷選手の第3打席内野ゴロ @ABEMA で視聴中
	Literal Translation	[Breaking news] Ohtani's third at bat was a groundout to infield watching on @ABEMA
	Japanese	栄冠ナインアンチ「地方大会で150km/hオーバーの高校生がゴロゴロ出るのはおかしい」
	Literal Translation	Glory Nine Anti-"It's strange to see high school students over 150 km/h romping around in local tournaments."
Home run	Japanese	【エンゼルスVSヤンキース生中継】＃大谷 選手、第3打席は2ランホームラン！ 29号！！！！ ▼URLから無料視聴▼
	Literal Translation	[Angels vs Yankees Live] #Ohtani's 3rd at bat was 2-run home run! No. 29!!!! ▼ Watch for free at the URL below ▼
	Japanese	［速報］大谷選手の第4打席??イッター 32号！ @ABEMA で視聴中
	Literal Translation	[Breaking News] Ohtani's 4th at bat?? Gone No. 32! Watching on @ABEMA
	Japanese	本日は、ダブルヘッダー！大谷翔平100号ホームラン出るか？ @ABEMA で視聴中
	Literal Translation	Today is a doubleheader! Shohei Otani, will you hit your 100th home run? Watching on @ABEMA
Strikeout	Japanese	大谷二刀流DAY速報｜第３打席は変化球に空振り三振。@ABEMA で視聴中
	Literal Translation	Ohtani two way day breaking news ｜ Ohtani' 3rt at bat was a strikeout against a breaking ball. Watching on @ABEMA
	Japanese	［速報］大谷選手の第4打席三振 @ABEMA で視聴中
	Literal Translation	[Breaking news] Ohtani's 4th at bat was a strikeout watching on @ABEMA
	Japanese	大谷翔平選手の奪三振集！！ 4.2回80球4安打9奪三振1四球1失点＃大谷翔平 #MLB ＃野球 ＃エンゼルス
	Literal Translation	Shohei Ohtani's strikeouts collection!! 4.2 innings, 80 pitches, 4 hits, 9 strikeouts, 1 base on balls, 1 run #ShoheiOhtani #MLB #Baseball #Angels

… Correct tweet … Incorrect tweet

LLMs include GPT (Generative Pre-trained Transformer) (Radford et al., 2018). These models are called "foundation models" (Bommasani et al., 2021). A foundation model is any model that is trained on broad data that can be fine-tuned to a wide range of downstream tasks. However, there is an exact difference between the two. GPT is based on Transformer Decoder, and uses a left-to-right architecture, where every token can only attend to previous tokens in the self-attention layers of the Transformer. On the other hand, BERT is based on Transformer Encoder, and uses a Masked Language Model (MLM) architecture. MLM enables the representation to fuse the left and the right context, which allows us to pre-train a deep

Table 2: Dataset

Category	Number of tweets	Label
Flyout	100	0
Base on balls	100	1
Groundout	100	2
Home run	100	3
Strikeout	100	4

Table 3: Implementation environment

OS	Windows 10 pro
CPU	Intel(R) Core (TM) i9-11900
GPU	NVIDIA GeForce RTX 3080
RAM	128 GB
CUDA Version	11.1
Python Version	3.8.16
Pretrained Japanese BERT model	2.6 GB

Table 4: Precision, recall and f-measure in tweet classification

Category	Precision	Recall	F-measure
Flyout	1.00	0.76	0.86
Base on balls	0.99	0.79	0.88
Groundout	0.93	0.94	0.94
Home run	0.61	1.00	0.76
Strikeout	1.00	0.79	0.88

bidirectional Transformer. In addition to the MLM, BERT uses a Next Sentence Prediction (NSP) task that jointly pre-trains text-pair representations. GPT is good at sequence generation and summarization, while BERT is good at sequence classification and question answering. According to Devlin et al. (2018), the GLUE score of BERT is superior to that of GPT. In addition, GPT has a much larger number of parameters than BERT, which limits the number of PCs that can process it. Furthermore, GPT is expensive to fine-tune, so BERT was used in this research.

4. Versatility Verification

In Section 3, we performed fine-tuning using tweets to specialize BERT for tweet classification task, and then input tweets to the model for classification. In this Section, we examine its generality to inputs other than tweets.

4.1 Sports commentary

In this Section, we use sports commentary from commentators to keep track of the status of sports games in real-time as an alternative to Twitter. We use the BERT model constructed in Section 3, which is specialized for tweet classification task, to classify the sports commentary.

Sports commentary is a means for commentators to provide viewers with information about the sports game as needed while watching it. Since the commentators must watch the game in real time and communicate it in an easy-to-understand manner to the viewers, we believe that the sports commentary is highly reliable. In this research, we manually transcribed the highlights and digests of the games on YouTube, one of the OTT services, and input them into BERT for classification. In order to verify the versatility of the classification, 10 sports commentaries were used for classification in each category.

4.2 Results

The Precision, Recall, and F-measure for sports commentaries classified by BERT, which were fine-tuned to be specific to the tweet classification task, are shown in Table 6. As shown in Table 6, the F-measure exceeded 0.9 points in the four categories of "Flyout," "Base on balls," "Home run," and "Strikeout." These factors can be attributed to the fact that the sports commentaries are as straightforward about results as the tweets. These factors can be attributed to the fact that there were only five categories at this stage, and because the "Others" category had not been set up, almost all ambiguous tweets were categorized as "Home run."

5. Conclusion

In this research, we fine-tuned BERT using tweets to calculate its classification accuracy. In general, short and noisy texts such as tweets are considered difficult to understand context at the word-level with tokenization. However, the results show that tweets can be understood in context by fine-tuning LLMs such as BERT.

Reference

Bommasani, R., Hudson, D., Adeli, E., Altman, R., Arora, S., Arx, S., Bernstein, M., Bohg, J., Bosselut, A., Brunskill, E., et al. (2021). On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*.

Canadian Radio-television and Telecommunications Commission (2011). Results of the fact-finding exercise on the over-the-top programming services. http://www.crtc.gc.ca/eng/publications/reports/rp1110.pdf. (Accessed April 20, 2023)

Devlin, J., Chang, M., Lee, K., and Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of NAACL*.

Fujimoto, N. and Ushiama, T. (2021). Syouryou no jikkyou tweet karano jyoukyousuitei model to sports kansenshien [A model for situation estimation from a small amount of live tweets and support for sports spectators]. *DEIM2021*. (in Japanese)

Fujimoto, Y., Imai, R., Nakamura, K., Tanaka, S., Arima, N., and Arakawa, T. (2018). Research for reliability analysis concern-

Table 5: Examples of correct and incorrect tweets

Correct	Incorrect	Language	Tweets example
Flyout	Base on balls	Japanese	しかし中塚くんはセンターフライツーアウト三塁二塁のチャンス 盛岡第四 ものに出来ず！花巻東の中居くんのピッチングが光ます
		Literal Translation	However, Morioka-daiyon couldn't take advantage of two outs, runners on third and second by Nakatsuka-kun was flyout to center field! The pitching of Nakai-kun from Hanamaki-higashi shines.
Base on balls	Home run	Japanese	四番 田代くんはデッドボールツーアウト二塁 一塁ツーアウトながら先制 チャンス！バッターは相野くんツーアウトから先制して欲しい 相野くんフォアボールツーアウト満塁！この もらったチャンス 活かしたい花巻東！
		Literal Translation	No.4 Tashiro-kun was hit by a pitched ball, two outs, runners on second and first but still chance! The batter is Aino-kun. I want to get the first run from two outs. Aino-kun was a base on balls, two outs, the bases are loaded! Hanamaki-higashi wants to make the most of the opportunities they have been given.
Groundout	Home run	Japanese	佐藤くん セカンドゴロツーアウト一塁ツーアウトからの花巻東の機動力 足！ 佐藤くん！盗塁！！続くキャプテン黒沢くんは期待にこたえて同点タイムリー スリーベース！花巻東 同点に追いつくなおもツーアウト三塁！
		Literal Translation	Sato-kun was groundout to second, two outs, runner on first. Hanamaki-higashi's mobility legs! Sato-kun! Steal a base!! Next, Kurosawa-kun, the captain, met expectations with game-tying RBI triple! Hanamaki-higashi equalized the game and still two outs, runner on third!
		Japanese	何とか最後はサードゴロチェンジ同点で止めたか花巻東VS八工大一4回表に 同点に追いつかれ2対2！ひとつのプレーから崩れる野球は怖いさぁ！花巻東！ しっかり粘ったバッティングで勝ち越しを！
		Literal Translation	Hanamaki-higashi vs Hachikou-daiichi, got a groundout to third and managed to stop the run tied, and the inning changed. The score was tied in the top the fourth inning, 2-2! Baseball that breaks down from one play is scary. C'mon! Hanamaki-higashi! Please overtake the game by steady and persistent batting!
Strikeout	Home run	Japanese	10勝目を狙う 大谷翔平初回 圧巻のワンマンショー 3連打でノーアウト満塁 から大谷翔平は即修正！ここからが大谷翔平のショータイム開幕三者連続三振 素晴らしい＃大谷翔平
		Literal Translation	Shohei Ohtani, aiming for his 10th win, puts on an impressive one-man show in the first inning. After three streaks, no outs, the bases are loaded, Shohei Ohtani immediately corrected himself! The real showtime begins for Shohei Ohtani, he got three strikes out in a row, awesome #ShoheiOhtani
		Japanese	小澤くん 三振結局 花巻東 3回表も無得点やはり打線は水物なのか？どこか チグハグどこかで県大会に爆発した打撃が目覚めるか！？ 花巻東！！
		Literal Translation	Kozawa-kun was strikeout after all. Hanamaki-higashi got no runs in the top of the third inning again. As expected, is an offence a game of chance? Somewhat mismatch, at some point, the offence that exploded in the prefectural tournament will awaken!? Hanamaki-higashi!!

Table 6: Precision, Recall and F-measure in sports commentary classification

Category	Precision	Recall	F-measure
Flyout	0.83	1.00	0.91
Base on balls	0.90	0.90	0.90
Groundout	0.89	0.80	0.84
Home run	0.90	0.90	0.90
Strikeout	1.00	0.90	0.95

ing microbloggers posting actual situations of road use to SNS. *IPSJ Journal*, Vol. 59, No. 10, 1853-1865.

Ministry of Economy, Trade and Industry and Japan Sports Agency (2022). Sports DX report. https://www.meti.go.jp/shingikai/mono_info_service/sports_content/pdf/20221207_1.pdf. (Accessed April 20, 2023)

MLB Advanced Media (2023). Pitch timer (2023 rule change). https://www.mlb.com/glossary/rules/pitch-timer. (Accessed April 20, 2023)

Ohmi, T., Kanada, K., Morinaga, M., and Emami, A. (2021). BERT niyoru shizen gengo shori nyuumon: Transformers wo tsukatta jissen programming [Introduction to natural language processing with BERT: Practical programming using transformers]. Ohmsha, Vol. 1, No. 4. (in Japanese)

Radford, A., Narasimhan, K., Salimans, T., and Sutskever, I. (2018). Improving language understanding by generative pre-training. https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf. (Accessed May 26, 2023)

Sakamoto, K., Nakamura, K., Yamamoto Y., and Tanaka, S. (2018). Research for social sensing technology to different phenomena during normal times. *IPSJ Journal*, Vol. 59, No. 10, 1866-1879.

Sasaki, T., Okazaki, M., and Matsuo, Y. (2010). Earthquake shakes Twitter users: Real-time event detection by social sensors. *Proceedings of the 19th International Conference on World Wide Web*. ACM.

Sasaki, T. and Matsuo, Y. (2012). Twitter as a social sensor: Can social sensors exceed physical sensors? *Journal of the Japanese Society for Artificial Intelligence*, Vol. 27, No. 1, 67-74.

Statista (2022). Average length of major league baseball games from 2000 to 2022. https://www.statista.com/statistics/1310998/mlb-game-length/. (Accessed April 20, 2023)

Stockmark (2022). Bert-book. https://github.com/stockmark-team/bert-book. (Accessed May 26, 2023)

The Hoops Geek (2022). How long is an NBA game?: Average NBA game length. https://www.thehoopsgeek.com/average-nba-game-length/. (Accessed April 20, 2023)

The Office of the Commissioner of Baseball (2023). Official baseball rules 2023 edition. https://img.mlbstatic.com/mlb-images/image/upload/mlb/wqn5ah4c3qtivwx3jatm.pdf. (Accessed May 26, 2023)

Tohoku NLP Group (2019). Pretrained Japanese BERT models released. https://www.nlp.ecei.tohoku.ac.jp/news-release/3284/. (Accessed April 20, 2023)

Toriumi, F. and Baba, S. (2016). Real-time tweet classification in disaster situation. *Proceedings of the 25th International Conference Companion on World Wide Web - WWW '16 Companion. the 25th International Conference Companion*.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, L., and Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 6000-6010.

X Corp. (2023). About. https://about.twitter.com/en. (Accessed May 26, 2023)

X Corp. (2023). Twitter privacy policy. https://twitter.com/en/privacy. (Accessed May 26, 2023)

(Received: May 2, 2023; Accepted: June 21, 2023)