

# AI を用いたスポーツに関するインターネットコメント のセンチメント分析方法の検討

河野 洋<sup>1)</sup>・藤本 太陽<sup>1)</sup>・北川 純也<sup>2)</sup>・河野 夏美<sup>2)</sup>

1) 福山平成大学 (福祉健康学部健康スポーツ科学科)

2) 順天堂大学 (スポーツ健康科学部)

E-mail : kohno@heisei-u.ac.jp

## 【要旨】

本研究は、スポーツに関するインターネットコメントのセンチメント分析に AI を用いた際に生じる、AI と作業者との評価の相違に関する知見の獲得を目的とした。本研究では、スポーツに関するコメントの評価では考慮すべき「スポーツの文脈」が存在するという仮説を立て、AI と作業者との相違を「スポーツの文脈」の影響の点から検証した。

調査として、「Yahoo! ニュース」に掲載された 2022 年の FIFA ワールドカップに関するニュース記事へのコメントから 1,000 件をランダムにサンプリングし、AI と 3 名の作業者がセンチメント分析をそれぞれ実施した。本研究の調査では、AI のセンチメント分析に OpenAI の “text-davinci-003” を使用した。

結果として、3 名の作業者が個別に実施したセンチメント分析では、全体の 51.1 パーセントのコメントで全員の評価が一致した。また、作業者全員の評価が一致したコメントを対象に AI と作業者とのセンチメント分析の結果を比較したところ、ポジティブは 93.0 パーセント、ニュートラルは 47.0 パーセント、ネガティブは 80.5 パーセントで両者の評価が一致した。

本研究の調査では AI と作業者とのコメント評価の相違として、① AI が「感情的でない」と評価したコメントを作業者が「感情的である」と評価する、② AI がネガティブと評価したコメントを作業者がニュートラルと評価する、という傾向がみられた。この傾向は、作業者がコメントを「スポーツの文脈」で評価することによって生じたと考えられる。

キーワード：インターネットコメント，センチメント分析，GPT，スポーツの文脈

## 1. 緒言

### 1.(1) センチメント分析の現状と課題

センチメント分析 (sentiment analysis) は、テキストや音声等のデータを基に、人の感情を分析する手法である。センチメント分析の一般的な手法としては、表出する感情に応じてデータを「ポジティブ」「ニュートラル」「ネガティブ」に分類するものがある。

今日、センチメント分析は社会の中においても多くの利用事例がある。企業は自社のサービス向上のため、顧客の声をポジ／ネガに分類し、サービス改善の知見として活用している。また、サービスや製品について多くの人がポジティブと認識する単語が何かを明らかにし、キャッチコピーに使用する事例等がある。

研究領域においても、センチメント分析を実施した事例は数多く存在する。特に、近年センチメント分析が実施されている分野のひとつが、インターネットやSNS上のコメント分析である。難波・福田<sup>1)</sup>は、新型コロナウイルスワクチンに関するX (ツイッター) の投稿から、人々のワクチンに対する「不安」の抽出を行った。また、瀧川<sup>2)</sup>は元首相の安倍晋三氏を対象に、世論研究におけるソーシャルメディアの言論分析の有用性を検証した。そこでは、安倍氏に対する投稿のポジ／ネガの分類と、世論調査による内閣不支持率とが比較されている。スポーツ領域においては、河野・北村<sup>3)</sup>がFIFAワールドカップを対象に行った「ウェブ上のレイシズム」の調査で、「差別的なコメント」の抽出にセンチメント分析の手法を用いている。

このようなセンチメント分析であるが、課題も存在する。テキストを扱うセンチメント分析の方法のひとつは、単語や表現と感情とを紐づけた「辞書」を作成し、辞書に従って分類を行う方法である。この方法は結果の客観性や再現性を担保するが、辞書の精度や妥当性が求められる。また、文脈によって単語や表現のニュアンスが変わる場合、機械的な作業が読み手の認識との間に齟齬を生じさせる恐れがある。たとえば、河野・北村<sup>3)</sup>はコーディング作業の中で「皮肉」を扱っているが、「気を付けて帰国してください」というコメントは、敗戦時に投稿されれば「帰国後にファンに襲われないよう気を付けて」という皮肉になる。

テキストを扱うセンチメント分析では、一人または複数の作業員によってデータを精査し、探索的に分類を行う方法が取られることもある。この方法では実際に読み手の立場になって分析を行うことができ、文脈による

ニュアンスの違いや皮肉のような単語だけでは判断できない感情の表出を考慮した分類が可能となる。一方で、感情を評価する基準を作業員に委ねる方法は客観性の担保が難しく、あるテキストの分類が「なぜそうなったのか」を説明することが困難な場合がある。さらに、作業員自身が評価基準を明確にできておらず、その都度感覚的な分類を繰り返した場合、結果の信頼性は大きく損なわれることとなる。

### 1.(2) AI を用いたセンチメント分析

近年は、センチメント分析にAIを用いた事例が数多く報告されるようになってきている。AIの進化については別稿に譲るが、研究の関心としてもAIによって人の感情を解釈しようとする試みは主要なテーマのひとつとなっている。

AIによるセンチメント分析は、少なくとも2点において有用性が認められる。1点目は、作業員の中で生じる評価の「ぶれ」をAIが最大限に排除できることである。あるデータのまとまりについて、作業員が最初から最後までを同じ判断で分類することは容易ではない。作業に時間的間隔を空けたり、似た内容のデータを繰り返し見たりすることによって、先ほどまでポジティブに分類していた内容のデータを途中からニュートラルに分類するといった、評価の「ぶれ」が起きることがある。これに対しAIは、分析作業で作業員に対して生じうる種々の影響を受けることはない。そのため、分類結果には一定の一貫性を認めることができる。2点目は、人には扱いが困難な大量のデータをAIが扱うことができることである。先述のインターネットやSNSのコメント分析の場合、日々投稿される大量のデータを処理できることが求められる。それを実現する方策として、分析にAIを用いることは有用なことだといえる。

以上のことから、人の感情に着目する際も、AIによるセンチメント分析の有用性を認識し、分析作業にAIを取り入れることについて検討することが必要であると考える。

### 1.(3) 研究の関心

本研究では、近年事例が報告されているスポーツに関するインターネットコメントの分析に、AIを取り入れることについて検討を行うこととした。特に、インターネットコメントに関する問題意識のひとつに、アスリート等への誹謗中傷や炎上がある。これらの投稿はセンチ

メント分析でネガティブに分類されるため、大量のコメントを一貫して分類できるAI分析にメリットがあると考えられる。

AIによるセンチメント分析を研究に取り入れる際は、AIによる分析結果がどのような傾向や特徴を持つかを把握し、データの扱いの指針を事前に決めておくことが重要だと考えられる。今日におけるAIのデータの扱いとしては、そのデータを無条件に信頼するよりも、樋口<sup>4)</sup>の「計量テキスト分析」が提唱するような、AIと作業者との批判的な相互作業によってその有用性が認められる。

AIによるセンチメント分析を研究利用する際、AIと作業者との分析結果の相違をどのように解釈するかは課題となる。解釈のひとつは、先述した作業者による評価の「ぶれ」である。この場合、AIの結果は作業者のぶれを是正するものとなり、両者の分析結果が一致することは分類の妥当性や信頼性を保証する根拠となる。

一方、文脈を考慮したセンチメント分析（contextual sentiment analysis）において、スポーツや政治、教育といったテーマはデータの解釈に影響を与える文脈として認識される。もし、AIによるセンチメント分析がスポーツという特定のテーマにとらわれない「標準的」なものであるとすれば、作業者は読み手の視点に立って分析をする中で、そのコメントが「スポーツの文脈」で投稿されていることを考慮するであろう。たとえば、政治というテーマが理性的なコミュニケーションを要求するものだとすれば、スポーツというテーマはユーザの感情的なコミュニケーションを許容するものと考えられる。水野ら<sup>5)</sup>が報告するプロ野球の試合内容に反応して生じるツイート・リツイートの「バースト現象」は、インターネット上における「熱狂」のコミュニケーションと捉えることができる。また、プーイングやトラッシュトーク、阪神ファンにみられる「アンチ巨人傾向」<sup>6)</sup>などは、対戦相手への敵対的・攻撃的態度がスポーツ文化として許容されていることを示すものといえる。こういった「スポーツであれば」という文脈が考慮される場合、AIと作業者との相違は単に是正されるべきものではなく、スポーツという特定のテーマで分析を行う際の重要な知見となる可能性がある。

ただし、AIと作業者との相違の理解については、相違がどの程度生じるか、そもそも作業者同士で相違が生じるのかどうかなど、実際に作業を行った上での検証が必要となる。

#### 1.(4) 研究の目的

本研究は、スポーツに関するインターネットコメントの調査における、AIを用いたセンチメント分析の実施に求められる知見の獲得を目的とした。

本研究では、以下の3つのリサーチ・クエスチョンを設定した。

- i. スポーツに関するインターネットコメントについて、複数の作業者が個別にセンチメント分析を行った場合の結果の一致度はどの程度か
- ii. AIによるセンチメント分析と、作業者による分析との結果の一致度はどの程度か
- iii. AIと作業者との分析結果の相違に、「スポーツの文脈」の影響が認められるか

なお、本研究では2022年にカタールで開催されたFIFAワールドカップ（以降、「W杯」とする）での日本代表に関するインターネットコメントをデータとして使用することとした。当該事例については大会期間中にマスメディアやインターネットメディアで連日報道がなされ、センチメント分析の作業者が大会内容やチームの動向等について一定の情報を持つことが期待された。一方で、サッカーという競技やW杯について、作業者が高度な専門知識を有する必要がない点も本研究の調査データとして適切であると考えた。

また、本研究で使用するAIについては、“OpenAI”が提供する“GPT”（Generative Pretrained Transformer）を採用した。GPTは昨今注目されているサービス“ChatGPT”の言語モデルとして使用されているものであり、一般的な利用実績や導入の容易さ、研究利用の汎用性等から判断した。

## 2. 方法

### 2.(1) データの収集・選定

インターネットニュース配信サービス「Yahoo! ニュース」（以降、「Yニュース」とする）でW杯期間中（2022年11月20日～12月18日）に掲載された大会に関するニュースを収集した。また、Yニュースには記事に対しユーザがコメントを書き込める機能があるため、ニュースに対するコメントもテキストデータとして収集した。

続いて、収集されたすべてのニュース記事より、本文を精査し日本代表の試合に関する記事のみを選定した。

## 2.(2) 分析用データのサンプリング

選定されたニュース記事に連なるコメント全体から、1,000件のコメントをランダムにサンプリングし、分析用データセットを作成した。

## 2.(3) AIによるセンチメント分析の実施

分析用データセットのコメントについて、AIによるセンチメント分析の結果を取得した。分析には“GPT-3.5”系のモデルのひとつである“text-davinci-003”のAPIを使用した。

具体的な作業としては、「Google スプレッドシート」のアドオン“GPT for Sheets™ and Docs™”が提供する関数“GPT\_CLASSIFY”を使用し、コメントを“positive”“neutral”“negative”のいずれか（または“ERROR”）に分類した。

表1 センチメント分析の結果

	ポジティブ		ニュートラル		ネガティブ	
	件数	割合	件数	割合	件数	割合
作業員A	363	36.3%	305	30.5%	331	33.1%
作業員B	297	29.7%	476	47.6%	226	22.6%
作業員C	284	28.4%	306	30.6%	409	40.9%
AI	243	24.3%	572	57.3%	184	18.4%

n=999

## 2.(4) 作業員によるセンチメント分析の実施

分析用データセットのコメントについて、3名の作業員が個別にセンチメント分析を実施した。作業員はコメントの内容を1件ずつ精査し、「ポジティブ」「ニュートラル」「ネガティブ」のいずれかに分類をした。作業員はいずれもスポーツ領域における研究者（スポーツ心理学者1名、スポーツマネジメント研究者2名）であり、スポーツ活動やスポーツ文化の知識を有する者であった。また、日常的にインターネットコメントを閲覧する習慣のない者であった。

## 2.(5) データの集計・比較

作業員のセンチメント分析の結果は、作業員ごとにポジティブ・ニュートラル・ネガティブの割合を算出した。また、コメントを、①3名の作業員すべての分類が一致したもの、②3名のうちの2名の作業員の分類が一致したもの、③3名の分類がすべて異なっていたものに分類し、それぞれの割合を算出した。

AIによるセンチメント分析の結果も、作業員の結果と同様にポジティブ・ニュートラル・ネガティブの割合を算出した。

次に、上記①に分類されたコメントを対象に、AIによるセンチメント分析の結果と作業員の結果との比較を行った。比較はAIの結果を基準とし、たとえばAIがポジティブに分類したコメントに対し、作業員がポジティブ・ニュートラル・ネガティブにそれぞれどの程度分類をしているかを割合で示した。

表2 作業員間のコメント分類の一致度

	ポジティブ	ニュートラル	ネガティブ	合計	割合
3人が一致	214	143	153	510	51.1%
2人が一致	105	194	167	466	46.6%
一致せず	23			23	2.3%

n=999

## 3. 結果

### 3.(1) 最終的な分析対象データの数について

本研究の調査では1,000件の分析用データセットを作成したが、AIの分析の際に1件のコメントがエラーとなったため、実際に分析対象となったデータは999件であった。

### 3.(2) 作業員およびAIのコメント分類の内訳

作業員のセンチメント分析の結果について、作業員Aはポジティブが36.3パーセント、作業員Bはニュートラルが47.6パーセント、作業員Cはネガティブが40.9パーセントとなり、それぞれ最も高い割合となった。

AIの分析結果については、ニュートラルの割合が57.3パーセントと最も高く、次いでポジティブ、ネガティブの順に割合が高かった。割合の傾向としては、作業員Bが同様の結果を示した（表1）。

### 3.(3) 作業員間のセンチメント分析結果の一致度

3名の作業員のセンチメント分析結果の一致度については、999件のうちの510件のコメントで、3名の分類が一致した（表2）。

### 3.(4) AIと作業員とのコメント分類の比較

作業員の分類が一致した510件のコメントについて、AIによるセンチメント分析の結果との比較を行った。

両者の結果の一致度について、AIがポジティブに分類したコメントのうち、作業員がポジティブに分類したコメントの割合は93.0パーセントであった。同様に、ニュートラルでの一致の割合は47.0パーセント、ネガティブでの一致の割合は80.5パーセントであった。

AIと作業者との結果に相違が見られた点としては、AIがニュートラルに分類したコメントについて、作業者がネガティブに分類するケース(32.3%)やポジティブに分類するケース(20.7%)があった。また、AIがネガティブに分類したコメントを作業者がニュートラルに分類するケース(17.2%)が認められた(表3)。

#### 4. 考察

##### 4.(1) 作業者のセンチメント分析の結果について

本研究の調査では、3名の作業者がW杯に関するインターネットコメントのセンチメント分析を行った。各々の分析結果については、最も高い割合で出現した感情評価が3名ですべて異なるという結果になった。

この結果は第一に、インターネットコメントに対する読み手の解釈に大きな違いがあることを意味するものといえる。実際のところ、この結果はインターネットコメントに関する研究に困難を与えるものとなる。探索的なセンチメント分析は読み手の解釈により近づくようとする試みであったが、それによって得られる結果はインターネットユーザのごく一部を代表するものとしかならないこととなる。センチメント分析にあたっては、扱うデータや想定する読み手の特性を念頭に、どのような作業者を選定するかを考える必要があるといえる。

ただし、今回の結果の意味するところは、単に「コメントの評価は人それぞれ」であるとか、「どのような

コメントもネガティブに捉える性格の人がいる」ということではない。今回の調査において、分析対象となったコメントの約半数は3名の作業者の分類が一致した。このように、インターネットコメントについては「多くの人の間で解釈が一致するもの」と「人によって解釈が異なるもの」があると考えられる(表4)。研究の関心としては後者のようなコメントの存在について、どのようなトピックや単語・表現が人々の解釈に相違を生じさせるか、またどのようなパーソナリティが感情の評価にどのような傾向をもたらすか、明らかにされることも必要だと考えられる。

##### 4.(2) AIおよび作業者のセンチメント分析の一致度について

本研究の調査において、3名の作業者の間で分類が一致したコメントのみをAIとの比較に用いた。この3名の分類が一致したコメントは、評価の「ぶれ」を排除し、多くの人の間でコメントの解釈が一致する見込みのあるものとなる。

AIと作業者との結果の比較においては、ポジティブで約90パーセント、ネガティブで約80パーセントのコメント評価が一致した。この割合について直ちに高い/低いの評価はできないが、場合によってはポジティブの一致度の高さは、作業者の分類の「ぶれ」の小ささを示す根拠となる可能性がある。

表3 AIと作業者との分類結果の比較

		作業者の結果						総計
		ポジティブ		ニュートラル		ネガティブ		
		件数	割合	件数	割合	件数	割合	
AIの結果	ポジティブ	160	93.0%	10	5.8%	2	1.2%	172
	ニュートラル	52	20.7%	118	47.0%	81	32.3%	251
	ネガティブ	2	2.3%	15	17.2%	70	80.5%	87

n=510

表4 コメントと作業者の分類結果の一例

作業者の分類結果	コメント
全員の分類が一致	悔しいけど、見ていて本当に応援したくなるチームだった。今後が楽しみ。これからも応援したい。(ポジティブ)
	とてつもないプレッシャーの中で蹴るPKは、このピッチに立った選手しかわからないんだろなあ。(ニュートラル)
	退屈だった。大河を見るべきだった。(ネガティブ)
全員の分類が異なる	まさに手のひら返しという言葉がぴったりの大会でしたね。
	失敗と書かず、止められたとかにしてよ
	ドイツがワールドカップで日韓に負けるなんて、誰も思えない。

表5 AIと作業者との分類が一致しなかったコメントの例

AIの分類	作業者の分類	コメント本文
ニュートラル	ポジティブ	正直負けると思っていました。総力戦でよく戦い抜きました！
		流行語大賞は村神様だけど 今日堂安様～
		ン、残念!でも、毎日日本戦が楽しみな数週間でした。選手の皆さん、ありがとう!
	ネガティブ	草サッカーを見ているようだった。
		ビッグマウスもいいが、結果出してから行ってくれ プレッシャーに負けてへなちょこキックだったな みんな
ネガティブ	ニュートラル	そうね、最近の日本の失点と敗戦パターンになってるもん。
		これがW杯。いくらい試合して押ししていたとしても、一瞬でも緩んだプレーしたらすぐ狙われる。相手も死ぬ気で来ている。
		勝ってほしかった。PK戦での負けはしょうがないと言うけどやはり悔しい。今日ももう寝れそうにないな。

一方、AIがニュートラルと分類したコメントのうち、作業者もニュートラルに分類したコメントは47パーセントと、ポジティブ・ネガティブに比べて一致度が低い結果となった。この結果は、AIが『感情的でない』と評価したコメントに対しても、作業者がポジティブやネガティブの感情表出を認める傾向にあったことを意味するものである。

#### 4.(3) 分類結果の相違に対する「スポーツの文脈」からの検討

AIと作業者との分類結果に相違が生じた原因として、本研究の関心である「スポーツの文脈」からの考察が可能であると考えられる。今回の結果は、一般的には感情的でないと考えられるコメント内容やコメント中の表現が、スポーツの文脈では感情的なものとして評価される可能性があることを示唆している。これは書き手がニュートラルな感情で投稿したコメントが、読み手によってポジティブにも、ネガティブにも意味付けられることとなる。これは、今回扱ったW杯というテーマや試合結果等がコメントを感情的なものとして捉えさせたミクロの可能性と、スポーツそれ自体がコメントを感情的なものとして解釈させるマクロの可能性の両方が考えられる。

本研究の結果の中では、AIがネガティブと判断したコメントを作業者がニュートラルに分類する傾向も認めることができる。スポーツを楽しむ中では、怒りや悔しさといった感情が生じたり、敵と味方に分かれる構図が相手への攻撃的な態度として現れたりすることがある。ただし、それらの感情は純粋な憎悪や攻撃性としてではなく、一部はスポーツでの自由な感情表

現として許容されることもある。今回の結果の場合、一般的にはネガティブと評価される単語や表現について、作業者がそれをスポーツの文脈においては許容されるニュートラルなものとして評価したと考えられる。

このスポーツの文脈を考慮した解釈はスポーツに関するインターネットコメントの分析における重要な要素となる。たとえば、スポーツというテーマ特有のコミュニケーションの中で「ネガティブな意図はない」という同意が得られている単語や表現が、AIの標準的な分析によってネガティブと評価されるようなことが起こりうる。ユーザにとってネガティブでないものをAIがネガティブと評価することは、場合によっては誹謗中傷や炎上のリスクを過大評価するものとなる。一方で、この結果はスポーツの文脈において、誹謗中傷や炎上といったネガティブなコメントが許容される危険性を示すものとも捉えることができる。その場合はAIの結果が、過度にネガティブなコメントを許容する恐れのあるスポーツの文脈を是正するものとなる。

#### 5. まとめ

本研究の目的は、スポーツに関するインターネットコメントの調査における、AIを用いたセンチメント分析の実施に求められる知見を獲得することであった。「Yahoo! ニュース」に掲載された2022年のFIFAワールドカップに関するコメントについて、AIと3名の作業者がそれぞれセンチメント分析を行い、比較をした。結果として、3名の作業者が個別に実施したセンチメント分析では、全体の51.1パーセントのコメントで全員の評価が一致した。また、作業者全員の評価が一致したコメントを対象にAIと作業者とのセンチメント

分析の結果を比較したところ、ポジティブは93.0パーセント、ニュートラルは47.0パーセント、ネガティブは80.5パーセントで両者の評価が一致した。本研究の調査ではAIと作業者とのコメント評価の相違として、①AIが「感情的でない」と評価したコメントを作業者が「感情的である」と評価する、②AIがネガティブと評価したコメントを作業者がニュートラルと評価する、という傾向がみられた。

本研究の結果については、AIと作業者とのセンチメント分析結果の相違をもたらす「スポーツの文脈」の影響が示唆された。このことは、大勢の読み手によってコメントの評価が一致した場合、それが必ずしもAIの評価に近づくわけではないことを意味する。このスポーツに関心を持つインターネットユーザの感情評価の基準は、今後の研究の関心のひとつとなる。一方で、読み手の解釈の相違は、スポーツを享受する者を分類するクラスターを想起させるものでもある。AIと読み手との相違と同時に、読み手の間にある相違も、今後関心を持って調査される必要がある。

本研究で扱ったW杯という事例は、スポーツに関するインターネット上の感情表出のごく一部を扱っているに過ぎない。今後は他の事例でも同様の調査を行い、「ポジティブ・ネガティブはAIと作業者との評価の一致度が高い」「ニュートラルは一致度が低い」という結果について検証を重ねることも必要であると考えられる。

#### 参考文献

- 1) 難波英嗣, 福田悟志. (2021). ネットからの不安感の情報抽出. 感性工学, 19(4), 163-170.
- 2) 瀧川裕貴, 永吉希久子, 呂沢宇, 下窪拓也, 渡辺誓司, 中村美子. (2023). ソーシャルメディア言論分析の方法①. 放送研究と調査, 73(3), 70-85.
- 3) 河野洋, 北村薫. (2017). スポーツの国際大会とウェブ上のレイシズム: 2014FIFAワールドカップブラジルに関する日本のウェブコメントに着目して. スポーツ産業学研究, 27(2), 149-162.
- 4) 樋口耕一. (2020). 社会調査のための計量テキスト分析 (2nd ed.). ナカニシヤ出版.
- 5) 水野誠, 佐野幸恵, 笹原和俊. (2021). 熱狂するファンダム—プロ野球ファンのツイートを分析する—. マーケティングジャーナル, 40(4), 6-18.
- 6) 広沢俊宗, 小城英子. (2005). プロ野球ファンに関する研究 (I): 阪神ファンと巨人ファンの比較 (第一部 地域と生活). 関西国際大学地域研究所叢書, 2, 3-18.

# A Review of Sentiment Analysis Method of Internet Comments About Sports Using AI

Yoh KOHNO<sup>1)</sup>, Taiyo FUJIMOTO<sup>1)</sup>, Junya KITAGAWA<sup>2)</sup>, Natsumi KOHNO<sup>2)</sup>

1) Fukuyama Heisei University  
2) Juntendo University

E-mail : kohno@heisei-u.ac.jp

## Abstract

The purpose of this study was to obtain knowledge about the differences in evaluation between AI and workers that occur when AI is used to sentiment analysis of Internet comments about sports. In this study, we hypothesized that there is a "sports context" that should be taken into account when evaluating comments about sports and examined the differences between AI and workers from the perspective of the influence of the "sports context".

As part of our research, we randomly sampled 1,000 comments from comments on news articles about the 2022 FIFA World Cup published on "Yahoo! News", and an AI and three workers each conducted sentiment analysis. In the research for this study, we used OpenAI's "text-davinci-003" for AI sentiment analysis.

As a result, in the sentiment analysis conducted by three workers individually, the evaluations of all the workers matched 51.1% of the total comments. Also, when comparing the results of sentiment analysis between the AI and the workers about the comments for which all workers' evaluations were the same, the evaluations of both sides were in agreement: 93.0% were positive, 47.0% were neutral, and 80.5% were negative.

In the research for this study, the differences in comment evaluation between AI and the workers are the tendencies that (1) comments that AI evaluates as unemotional are evaluated by the workers as emotional and (2) comments that AI evaluates as negative are evaluated by the workers as neutral. These tendencies are thought to be caused by the workers evaluating comments in a "sports context".

KEYWORDS : Internet comments, sentiment analysis, GPT, sports context