Analyzing Differences in Major League Baseball Interview Questions Asked to East Asian and American Players

Danny Kim

Abstract

Major League Baseball (MLB) has seen a significant rise in the relevance of foreign players from Asia in recent years, which has led to an increase in their coverage by sports media. I introduce a new dataset of MLB press conference interviews taken from ASAP Sports with labels indicating the nationality of the interviewee. To evaluate whether there is a clear difference between interview questions asked to American players and those asked to foreign players from Asia, I first extract contextual BERT embeddings on this data, then train a basic Logistic Regression classifier on the embeddings. The model exhibits a poor Recall score when classifying questions asked to East Asian players, suggesting that classification using the embeddings on this dataset does not reveal a clear widespread difference in the questions asked to East Asian players as opposed to American players. I cluster the embeddings using k-means to identify any outliers, finding that the clustering algorithm has difficulty forming distinct clusters. I further analyze these clusters in relation to the findings of the classification task, and perform qualitative analysis on examples from this data.

1 Introduction

Media coverage regarding East Asian baseball players has increased as players from Asia have grown in relevance and representation in Major League Baseball (MLB) over the years. Because more and more East Asian players are starting to represent their respective countries in MLB, the significance of representation and nationality in the media coverage of these athletes becomes a topic to examine. Doing so may reveal, for instance, that the media is concerned more about the player's nationality than their qualities as athletes. In particular, I want to examine whether or not East Asian baseball players are being asked similar interview questions to their American counterparts, and whether these differences are informed by the players' status as foreigners in America.

Because the questions asked by sports journalists in part reflect the interests of not only the journalist but the public audience of sports media, this research may reveal information relating to what the American public is curious about regarding foreign MLB players from Asia. This information may be used to determine whether foreign East Asian players, and by extension East Asian foreigners in general, are viewed differently by the American public, and evaluate the importance played on their identity in this perspective.

I assemble a dataset of MLB press conference questions and answers, along with metadata about the date and event of the interview as well as the player's name, taken from ASAP Sports. The metadata also includes the player's nationality, which is scraped from the players' Wikipedia entries. The dataset, as well as the code for assembling it, is made public for other researchers.[1]

In the preprocessing step for this research, I use a BERT model trained for named-entity recognition to censor the players' names, after which I extract contextual embeddings from the questions using BERT and perform a variety of analysis on these embeddings. I train and evaluate a basic Logistic Regression on the classifier on these embeddings that classifies whether a question is asked to an American player or an East Asian player. The model exhibits a poor Recall score for questions asked to East Asian players, suggesting that the questions asked to players belonging to each category (American, East Asian) are not uniformly different. I perform qualitative analysis on a subset of questions that the model correctly identifies as being directed to East Asian players, as well as the questions that the model false identifies as being

[1]The data and code can be found here: https://github.com/dtkim5/baseball-interview/

directed to East Asian players, in order to examine any commonalities between them.

I then perform clustering on the same embeddings in order to identify any notable outliers in these clusters. I find that the k-means algorithm does not form distinct clusters from these embeddings, suggesting that the questions are largely homogeneous in the context of these embeddings. Additionally, I examine the predictions from the classification task—the data that has been labeled as East Asian, correctly or otherwise—in relation to these clusters, to find potential patterns in the classification predictions. I find that while there are potential patterns to be identified from this examination, the sample size is too low and the clusters overall too similar to one another to confidently draw a conclusion.

Finally, I discuss the limitations of my data and research, and the future work that can be done on this topic and data.

2 Background

Established in the eponymous book by Edward Said, the concept of Orientalism (Said, 1978), in which the Western image of the Eastern identity is characterized as exotic and alien, is relevant in the analysis of sports interview questions directed at foreign East Asian players, which, as previously mentioned, can be viewed as an extension of the general American public's view of foreign East Asian identities. In particular, the Orientalist image of an exotic East may inform the fascination with East Asian players and their identities as such, that may be reflected in the differences this research attempts to identify.

Much of the recent surge in the media coverage of East Asian baseball players has been centered around Shohei Ohtani, a Japanese player with the unprecedented skillset of being able to both pitch and hit at an elite level. Because of his uniqueness as a player and his overwhelming success in Major League Baseball, Ohtani has become the face of baseball in America; as such, he has garnered much attention in sports media. Some of this attention is negative: in 2021, a prominent sports television personality was scrutinized for claiming that Ohtani was not "Marketable" in Major League Baseball because he speaks in interviews through an interpreter (The Athletic, 2021). In the same year, a commentator for the Detroit Tigers was suspended after speaking about Ohtani in a racist accent (Beck, 2021). These instances signify that the player's nationalities remains prominent in Americans' perceptions of them.

Previous research on using computational methods to identify bias in sports journalism has been conducted by Fu et al. (2016), which found that journalism in tennis exhibited widespread gender bias against female players, often asking them questions that did not pertain to the game of tennis and centering on the personal lives and appearance of the players. Merullo et al. (2019) investigated racial bias in American Football sports commentary, an undertaking which involved assembling a large corpus of American football commentary transcripts and manually-annotated labels for the racial identity of the players. Similarly, my research focuses on the impact of player identity on the language used around them, in an American sports context.

Sun et al. (2022) introduced a sports interview corpus, using transcripts taken from ASAP Sports and ESPN. This corpus, however, is centered around entity-centric dialogue, and has no identity labels for the interviewees. Additionally, while this corpus claims to release the data and the code to the public, no such data or code can be found. As such, no part of the dataset and code I introduce in this research is taken from their work.

3 Data

3.1 Web Scraping

The data for my research was taken from ASAP Sports[2], an online archive of interviews and press conferences in a variety of sports. For each sport, the website maintains a list of all interviewees and a list of all interviews that interviewee has participated in. Each interview page contains the event, date, names of the interview participants, and the raw transcript of the interview. The transcript of the interview is formatted as a series of questions and answers, the text of which are preceded by "Q." and "[PLAYER NAME]", respectively. An example is shown below:

> NL DIVISION SERIES: CARDINALS v DODGERS
> October 2, 2014
> Clayton Kershaw
> ...
> Q.Â Can you talk about[...]the goal of winning a World Series in Los Angeles

[2] https://www.asapsports.com/

for the first time since 1988.
CLAYTON KERSHAW:Â I don't think we think about that.Â Any team the goal is to win a World Series[...]
Q.Â To clarify about my earlier question, [...]
CLAYTON KERSHAW:Â No.

The web scraping process is as such: For each letter of the alphabet, for each player whose name begins with that letter of the alphabet, and for each interview that player has appeared in: the metadata—the event, date, and interviewee name(s), is extracted from the top of the page and saved. The raw interview transcript is extracted and saved as well. Using the event information, the data is filtered such that only MLB events are saved. Afterwards, the raw text is processed. Any unusual special characters are filtered out of the text data, then raw interview transcript is separated into chunks of questions and answers; this is done by splitting the text with the delimiter "Q." Doing so allows for questions with multiple answers (e.g. questions directed at multiple players) to be captured as well. Afterwards, RegEx is used to detect the first instance of consecutive capital letters followed by a colon, e.g. "CLAYTON KERSHAW:". The question-answers chunk is separated into the question and the answers to that question. The answers are further split using the same RegEx outlined previously, such that a row of the dataset can be assembled with the question, answer, and metadata.

To identify the player name corresponding to each answer, and to avoid issues caused by OCR errors, which from the above example is shown to be quite frequent, the capitalized name is first made into lowercase, after which the Levenshtein distance between the name and each of the (also lowercased) names found in the metadata of the interview page is taken. The player whom the answer to the question belongs to is determined to be the name with the lowest Levenshtein distance. Additionally, I use a BERT model for Named-Entity Recognition[3] to detect any instances of names appearing in the questions, replacing each of them with the token <PN>; the output of this name replacement is saved separately.

3.2 Nationality Labels

The nationalities of the players were retrieved from their respective Wikipedia pages. The reasoning behind this decision is that a player's country of birth, which is commonplace information in sports databases, does not always represent the player's nationality or identity. For instance, a player may be born in one country but identify much more closely to another. One example is Rob Refsnyder, who was born in South Korea but was adopted by an American couple at five months old: his place of birth is listed as South Korea, but his experiences and perception in an American context are much more different than those of a native Korean who comes over to the MLB after living and playing in Korea for his entire life. As such, the player's Wikipedia entry, which may more accurately reflect the player's perceived nationality.

The Wikipedia API was used to get each article for the player. Because the archived interviews from ASAP Sports date back to as early as 1990, there are many inconsistencies in the formatting of each page. This results in a lot of other elements, such as the location the interview took plce in, part of the event, etc., leaking into the section of the page where the player names are listed. As such, some preliminary measures are taken to ensure that I am not accidentally retrieving the nationality of a city or baseball team instead of a player. Additionally, some non-players may be interviewed as part of a press-conference, e.g. doctors, team staff, family of the players. The Wikipedia API returns a list of the top N articles for the searched phrase; from each article that pops up, the summary can be retrieved. The summaries for Major League Baseball players are almost uniformly formatted as such:

> José Orlando Berríos (born May 27, 1994), nicknamed "La Makina" (Spanish for "The Machine"), is a **Puerto Rican** professional **baseball** pitcher for the Toronto Blue Jays of Major League Baseball (MLB). He previously played for the Minnesota Twins, who selected him in the first round of the 2012 Major League Baseball draft.

The relevant information from this summary has been bolded for visibility. First, the list of summaries is filtered for the word "baseball" appearing in it, taking the first summary that appears. Afterwards, a list of country names and their corresponding adjectivals (e.g. United States : American)is

[3] https://huggingface.co/dslim/distilbert-NER

retrieved from Wikipedia[4] and made into a dictionary, and the selected summary is searched for any of the adjectivals in the dictionary. As can be seen in the example, occasionally adjectivals such as "Spanish," that do not correspond to the player's nationality but rather the name's language of origin, may appear before the actual nationality in the summary. To circumvent this issue, the search only occurred after the first instance of a closed parenthesis. Altogether, this simple process filters for MLB players out of the messy list of names from the web scraping process, and retrieves their nationality. A dictionary that maps the player's name to his nationality is made from this information, which is used to attach the nationality label to the data during the original process of web scraping and processing the raw data into a dataset.

4 Methodology

The methodology of this paper is as such: BERT(Devlin et al., 2019) contextual embeddings are extracted from the questions and scaled. The labels of the East Asian countries represented in MLB: China, Japan, Korea, Taiwan, and Hong Kong, are aggregated into one label for the purpose of simplifying the analysis. A Logistic Regression classifier is trained on these embeddings to measure if the difference in the questions asked to American players versus East Asian players can be easily detected with a classifier. Afterwards, some distance-based analysis, involving clustering and qualitative analysis of classifier results using these clusters, is performed to make sense of the results in a more qualitative perspective.

5 Classification

A Logistic Regression classifier is trained on the scaled BERT contextual embeddings, with an 80-20 train-test split. This classifier is trained twice: first without the player names replaced, and then with the player names replaced. Without player name replacement, the results are shown below:

Table 1: Results of classifier - player names included.

Class	Precision	Recall	F-score	Support
East Asia	0.35	0.07	0.12	201
United States	0.99	1.00	0.99	15748

[4] https://en.wikipedia.org/wiki/List_of_adjectival_and_demonymic_forms_for_countries_and_nations

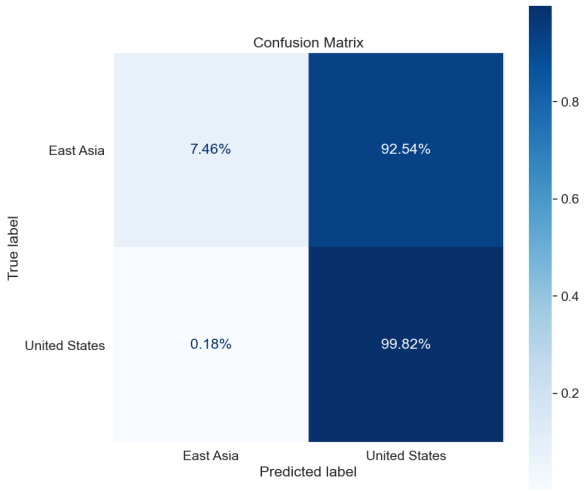

Figure 1: Confusion matrix of classifier - player names included.

As can be seen, the classifier exhibits a poor Recall score of 0.07, meaning that the classifier only correctly predicted 7% of questions directed towards East Asian players. This indicates that the differences in the questions asked to East Asian players versus American players is not uniformly detectable using these embeddings. In this section, I examine some questions directed to East Asian players and correctly labeled as such, and compare them to some questions directed to American players but incorrectly labeled as East Asian. Below are some examples of the former:

> Hiroki, you've thrown a career high in innings this year. How fresh does your arm feel, and have you felt the impact of those innings on your arm as the season wound down?
>
> Masahiro, when you look at your starts, what is the difference for you personally when you're sharp and good and when things get away from you?
>
> What differences have you found in the culture between here and Japan, and also the behavior of the fans?
>
> I just wanted to ask you: Can you tell us exactly how you feel now?

And below are some examples of the latter:

> I know you're facing Shohei at the WBC and now you're teammates with him. Is there something you learned, anything new, about Shohei?

> Is Yoshii still going today? There were some rumors around that he wasn't?
>
> What do you see in the two Japanese pitchers this year, Matsuzaka and Okajima?
>
> How did your arm feel when you made your throw?

Through the similarities in these examples, it can be posited that the model's criteria for a question being classified as East Asian includes: asking the player how he feels; mentions of East Asia-associated places such as Japan; and explicit mentions of the names of East Asian baseball players. The first two criteria may have a common link through the interviewer's question that asks an East Asian player how it feels to represent their home country in Major League Baseball, though this is loose and speculative. The explicit mention of names is the least useful of these criteria; they are most likely learned from the training data, in which interview questions, as they often do, begin by addressing the player by name. As such, I train and analyze the performance of the classifier on the name-replaced embeddings to see if replacing the player names would improve the model's performance. The results are shown below: It can be observed that

Table 2: Results of classifier - player names replaced.

Class	Precision	Recall	F-score	Support
East Asia	0.35	0.07	0.11	201
United States	0.99	1.00	0.99	15749

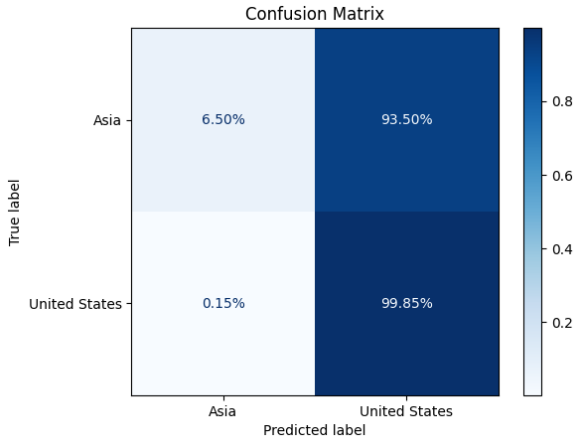

Figure 2: Confusion matrix of classifier - player names replaced.

the figures of the classification report are very similar to that of the classifier without player name replacement. Nevertheless some examples must be examined in a similar fashion as before, from the correctly labeled as East Asian category and the falsely labeled as East Asian category. Below are some examples of the former:

> Do you think it can help motivate you and maybe provide a little bit of extra energy knowing what an important game this is for your homeland and the number of people who will be watching, and how significant it is to have a South Korean pitcher start a Postseason game?
>
> I wonder if you could put into some perspective, it's very difficult for a Japanese playing to get here, play ten years in Japan and come to a new culture, and I'm sure you thought you'd win a championship before this. What does this journey feel like and how do you feel right now?
>
> What differences have you found in the culture between here and Japan, and also the behavior of the fans?
>
> I just wanted to ask you: Can you tell us exactly how you feel now?

And below are some examples of the latter:

> There's nine Venezuelans in this World Series. You managed in Venezuela yourself and you have managed many Venezuelan players. Could you talk about why you think it has evolved so much, baseball in Venezuela, and why there are so many good players in the majors?
>
> How many times do you pitch in elimination games in your career, college or minor leagues?
>
> When you see another team bench some of their powerful left-handed hitters against you, the stats indicate you're more effective against right-handers. How do you feel?
>
> And after the recovery you've been doing, how do you feel going into tomorrow?

It was observed that questions containing mentions of the East Asian players' names were no

longer a large portion, if present at all, of either of the categories labeled as East Asian. What has replaced these documents, in the category of correctly-labeled East Asian questions, are some more overt mentions of the player's nationality. The mention of how a player "feels" is still widely present in these documents, and becomes much more apparent in the questions falsely labeled as East Asian. A potential explanation for this is that a significant majority of East Asian players imported into Major League Baseball are pitchers; it is rarer for a hitter from the East Asian leagues to transfer his skills to be effective against MLB pitching. Pitchers are often asked about their own evaluation of their performance, which is often phrased as "how do you feel..." As such, the classifier may have picked up these features as indicators of nationality despite being unrelated to it. Overall, the replacement of player names resulted in more informative classification despite failing to increase the Recall score.

The consistently-low Recall score of the classifier trained on the embeddings of my data suggests that using these embeddings, it is difficult to consistently detect whether a question is directed to an East Asian player. However, the goal of this classification is not to achieve high classification metrics; the questions, regardless of the interviewee, for the most part are in the domain of baseball, and as such there is no guarantee that a given question asked to an East Asian player will be distinct from one asked to his American counterpart. The true goal is to use the predicted labels to determine the detectable aspects of these differences, and more generally determine whether such differences plausibly exist. The recall score is tremendously low such that it would be difficult to conclude that there is, generally, a difference in the questions asked to East Asian players and those asked to American players; however, the classification did show evidence of these differences occurring with some potential identifiable features between them. Analysis using distance-based metrics may allow for further analysis of these outliers as a whole rather than only those that fit a certain category according to the classifier, and may additionally make more apparent the causes of the classifier's difficulty with achieving high scores on the embeddings.

6 Distance-Based Analysis

Using the earlier scaled embeddings, k-means clustering is performed on the data, creating k=8 clusters (the final value of k is the value that returns the best silhouette average out of the values of k between 1 and 16). PCA is used to take the first two principal components of the embeddings so that the clusters can be visualized on two dimensions. As can be observed in Figure 3, K-means clustering does not form distinct clusters from these embeddings, suggesting that much of these interview questions are indeed homogenous. Nevertheless, each cluster is examined for its distribution of East Asian and American players represented. The cluster distribution for each category (American, East Asian), i.e. what questions in each category belong to which cluster, is additionally examined. It is difficult to determine anything from these figures alone, since there are no clear outliers in the clusters, i.e. clusters which have a significantly greater proportion of East Asian players than the rest. To relate the clusters to the findings of the classification task, I examine the cluster distributions of the two categories (true, false) of the questions labeled

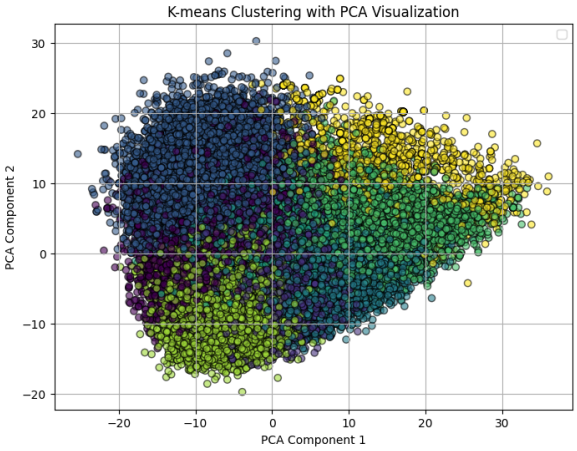

Figure 3: k-means clustering, with all 8 clusters plotted.

Cluster	American	East Asian
0	10227 (98.0%)	204 (2.0%)
1	4751 (99.4%)	29 (0.6%)
2	10751 (98.2%)	197 (1.8%)
3	4888 (99.2%)	40 (0.8%)
4	10884 (99.0%)	111 (1.0%)
5	10956 (98.4%)	179 (1.6%)
6	9009 (99.3%)	60 (0.7%)
7	1500 (99.5%)	8 (0.5%)
Total	62966	828

as East Asian by the classifier. For both categories,

Cluster	East Asian - True	East Asian - False
0	8 (61.5%)	8 (33.3%)
1	0	0
2	4 (30.8%)	4 (16.7%)
3	0	3 (12.5%)
4	0	4 (16.7%)
5	1 (0.08%)	2 (8.3%)
6	0	1 (4.2%)
7	0	2 (0.08%)
Total	13	24

the questions were clustered the most into clusters 0 and 2, in that order. Text from these clusters are examined to find any commonalities. Below are examples from Cluster 0 qualitatively deemed to be representative of the cluster subset:

> **East Asian - True:**
>
> *...knowing what an important game this is for your homeland [...] to have a South Korean pitcher start a Postseason game?*
>
> *...it's very difficult for a Japanese player to get here, play ten years in Japan and come to a new culture[...]*
>
> **East Asian - False:**
>
> *How deep do you think you'd be able to go if you're pitching well [...] how many pitches do you think you'll be able to throw?*
>
> *...What have you liked the most about pitching maybe at Comerica [...] What have you liked the most about Detroit?*

And from Cluster 2:

> **East Asian - True:**
>
> *...how do you feel about that? [...] how did you feel when you hit the home run?*
>
> *...You pitched great today for five innings. How did you feel about today's outing?*
>
> **East Asian - False:**
>
> *How does it feel to come back this season [...] How do you feel about everything?*
>
> *...your value [...] will go up or down dependent on how you pitch in postseason. How do you feel about that?*

Cluster 0 appears to contain questions that pertain to nationality for East Asian players, whereas for American players the cluster appears to contain questions pertaining to pitching and occasionally mentions of locations such as Detroit and Comerica Park. Cluster 2, for these categories, appears to contain questions in which the player is asked about how he feels, as well as, similar to Cluster 0, about pitching. However, it is difficult to determine whether Clusters 0 and 2 as a whole are actually centered around these topics; examining samples of the subset of data in these clusters that exclude the East Asian-labeled data reveal a wide variety of topics that often overlap between clusters. Additionally, as the number of the questions labeled as East Asian by the classifier is very low compared to the overall size of the dataset, it is difficult to confidently draw a conclusion about these commonalities.

7 Discussion

Although there are some outliers in the data that may indicate such, neither the classification or the clustering were able to identify strong evidence that East Asian players and American players are asked different questions. The classifier was only able to label a very small portion of the East Asian player interview questions as being directed to an East Asian player, and the clusters formed on the embeddings were difficult to distinguish from one another as a whole, though the clusters appear to potentially contain some distinct, feature-informed subcategories of the questions predicted to be directed to East Asian players by the classifier. In this section, I discuss the limitations of my research, and the ways in which future work could improve upon my research.

7.1 Limitations

One difficulty encountered in this research was that in the embedding space for this dataset, the datapoints were overall close together. This made it so that the classifier had difficulty distinguishing between the questions asked to American players and those asked to East Asian players, and the clustering algorithm had difficulty forming distinct clusters. One explanation for this is that because the base BERT model covers the general domain of English, the questions in the dataset, which are all, for the most part, about baseball, may be assigned similar positions in the embedding space even despite the scaling. A model adapted to the baseball domain may allow for differences between

the questions to be better captured.

This research is also limited the size and variety of the dataset. ASAP Sports only archives MLB press conferences for notable events such as the All-Star Game and the Postseason, which contributes to the relatively small number of East Asian baseball player interviews in the dataset. Additionally, this further narrows the scope of the dataset, as these events themselves are often the subject of the interview questions. Even in the case of questions pertaining to an East Asian player's nationality, many of them are given in the context of the playoffs; for instance, in one question in the dataset, pitcher Hyun-Jin Ryu is asked about the importance of a South Korean pitcher starting a Postseason game. As such, the topic of the questions are not neatly separated into nationality, pitching, playoffs, etc. but are amalgamations of a number of these topics, which is intensified by the additional context of an important event that is common to many of these questions. Had the dataset included interview questions for regular season games, the differences in the questions may have been easier to observe.

7.2 Future Work

Future work may address these limitations to improve upon my research. A BERT-based model may be fine-tuned on text relating to baseball—news articles, baseball commentary, online discussion of baseball—to better capture the finer differences in the baseball questions in its embeddings.

To improve the size and variety of the dataset, an approach similar to Sun et al. (2022) can be taken, in which more than one interview archive is aggregated into the dataset. The inclusion of interviews from regular season games would also allow interview questions asked to less-established East Asian players, such as those who are not offered interviews for postseason series, those who are not on teams that qualify for the postseason, and those who only have a short, troubled stint on a Major League roster before being reassigned to the Minor Leagues. These players may be the subjects of many interview questions regarding their nationality, as the interviewer may not know as much baseball-related information about the player as they may about a more prominent player.

Future work may also incorporate a subset of my data ignored in my research: Latin-American players. Latin-American players have a much longer history of success in Major League Baseball than players from East Asia; as such, comparing the questions asked to Latin-American players with those asked to East Asian players, potentially by comparing the distances of these questions from the centroid of the set of questions asked to American players, may further contextualize the findings of my research. Alternatively, the questions directed to Latin-American players may be combined with those directed to East Asian players as part of the set of non-American players to shift the focus of research to the differences in perception between American players and non-American players.

To answer the underlying question behind my research, regarding perceptions of East Asian identities in American sports, instead of using interview questions as proxy for national attitudes in America, sports social media data can be analyzed in future work to more directly examine these attitudes. This would increase the variety and size of the data and potentially allow for clearer patterns to be isolated.

8 Conclusion

The rise in the relevance of East Asian players in Major League Baseball (MLB) has led to increased media coverage of these players. In this paper, I introduce a dataset of baseball interview questions with metadata including the player's nationality. Using this data, I examine the interview questions asked to East Asian MLB players and compare them to the questions asked to their American counterparts. Our analysis fails to reveal a consistent trend of identifiable differences between these two categories, revealing the limitations of the data and embeddings. The data and code is publicly released in hopes of inspiring future work on the data or topic.

References

Jason Beck. 2021. Jack morris suspended for ohtani remarks. *MLB*.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Liye Fu, Cristian Danescu-Niculescu-Mizil, and Lillian Lee. 2016. Tie-breaker: Using language models to quantify gender bias in sports journalism.

Jack Merullo, Luke Yeh, Abram Handler, Alvin Grissom II, Brendan O'Connor, and Mohit Iyyer. 2019. Investigating sports commentator bias within a large corpus of American football broadcasts. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 6355–6361, Hong Kong, China. Association for Computational Linguistics.

E.W. Said. 1978. *Orientalism:*. Number v. 10 in Orientalism. Pantheon Books.

Hanfei Sun, Ziyuan Cao, and Diyi Yang. 2022. SPORTSINTERVIEW: A large-scale sports interview benchmark for entity-centric dialogues. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 5821–5828, Marseille, France. European Language Resources Association.

The Athletic, 2021. Espn's stephen a. smith apologizes for comments about shohei ohtani.