Presenting XAI-generated Explanations Of Cricket Shots

Gido Vitner[1]
Supervisors: Ujwal Gadiraju[1]**, Danning Zhan**[1]
[1]EEMCS, Delft University of Technology, The Netherlands

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
June 22, 2025

Name of the student: Gido Vitner
Final project course: CSE3000 Research Project
Thesis committee: Ujwal Gadiraju, Danning Zhan, Mark Neerincx

Abstract

Explainable Artificial Intelligence (XAI) has the potential to enhance user understanding and trust in AI systems, especially in domains where interpretability is crucial, such as cricket training. This study investigates the impact of different explanation formats on user experience within a cricket-specific context. Two prototypes were developed, each including four explanation formats: textual, visual, rule-based, and mixed. The second prototype introduced interactive features to examine their influence on user experience and explanation effectiveness. A small-scale user study evaluated the explanations based on satisfaction and trust. Results show that rule-based explanations were significantly less preferred in terms of satisfaction than the other explanation formats. Furthermore, the addition of interactive features led to a significant increase in user trust, though they did not enhance satisfaction levels. These findings highlight the importance of selecting appropriate explanation formats and the potential of interactive features to enhance trust in AI-generated explanations in a cricket-specific context.

1 Introduction

Cricket is one of the most popular sports around the world with tens of millions of amateur and professional players across all continents [15]. Like many other sports, cricket is increasingly benefiting from technological advancements aimed at improving player performance. Technological tools have shown promise in improving the learning process of a cricket player and could potentially improve cricket performance [13].

An example of such technology is the use of machine learning to automatically recognize and classify cricket shots from video footage or images [1]. This process can be further refined using techniques such as Human Pose Estimation [5, 11]. However, despite the growing accuracy of such models, understanding how and why a machine learning algorithm arrives at its decision remains a challenge, especially for end-users like cricket players who may not have a background in AI. This is where Explainable AI (XAI) becomes relevant.

XAI refers to a set of techniques that aim to make machine learning models more transparent, interpretable, and trustworthy. It does so by providing explanations for the outputs of these models, often by identifying the contribution of individual input features. Although XAI has the potential to improve the cricket learning process, the effectiveness of given explanations could depend on how they are presented. Explanations can vary widely in format, and not all are equally intuitive or useful, particularly for users not familiar with the field of machine learning.

This research investigates how XAI-generated explanations can best be presented to support learning in cricket training. The main research question is:

> What are the most effective ways to present XAI-generated explanations to facilitate learning in cricket training?

To address this question, the following subquestions are explored: 1) What types of explanation formats exist? 2) Which of these formats are most efficient for cricket training? 3) Do interactive features enhance the effectiveness of these explanations?

These subquestions are first addressed through a literature review in Section 2. Five explanation formats are identified, four of which are being implemented in prototypes simulating a cricket learning environment. A user study is then conducted to explore the second and third subquestions. The prototype and the format of the survey used in the user study are described in Section 3. In Section 4, the results of this study, including satisfaction and trust scores for each explanation in both a prototype with and without interactive explanations, as well as relevant statistical analyses. These findings are then interpreted and contextualized in Section 5, which also outlines the limitations of the prototype and user study. In Section 6, ethical aspects and reproducibility of the research are discussed. Finally, conclusions are drawn and directions for future research are proposed in Section 7.

2 Background

To address the first subquestion (*What types of explanation formats exist?*), the work of Vilone and Longo [14] provides a useful foundation. In their systematic literature review, aimed to organize output formats of XAI models, they identified four primary explanation formats: numeric, rule-based, textual, and visual. Additionally, they introduced a fifth category called mixed explanations, which combines elements from multiple formats.

In relation to the second subquestion (*Which of these formats are most efficient for cricket training?*), numeric explanations seem to be less suitable for general users, as they often lack intuitiveness and are harder to interpret without technical expertise.

Interestingly, studies in various domains suggest that the specific explanation format may not always significantly influence the perceived effectiveness of an explanation [2, 4, 16]. In these cases, the mere presence of an explanation mattered more than its format when measuring effectiveness. However, preference can vary based on specific contexts. For example, Kouki et al. [9, 10] examined explanation formats in the context of music recommender systems and found that textual explanations were preferred over visual ones when evaluating the persuasiveness of personalized explanations, though no such preference was observed when mock explanations were being used [8].

In a different domain, Girmay and Möhrle [6] investigated different explanation formats in the context of dairy farming systems. In their user study, they showed participants four different explanation formats and found no significant differences among them. Instead, they suggested that user preferences are highly context- and user-specific and that an ideal system would allow users to select from multiple explanation formats according to their needs.

Finally, the third subquestion (*Do interactive features enhance the effectiveness of these explanations?*) has also been explored in prior research. In the context of recommender systems, Kouki et al. [10] found no statistically significant difference between static and interactive explanations. However, broader literature suggests that interactivity may enhance user engagement and comprehension, especially in learning environments. For example, Bali et al. [3] studied the effects of interactive elements on the cognitive load of children. They showed groups of children different electronic books, some of which contained interactive elements, and concluded that interactive animated figures were more effective than static visuals in reducing cognitive load and supporting learning.

3 Methodology

Based on insights from existing literature, two prototypes were developed to explore the effectiveness of different explanation formats. This section outlines the design of these prototypes, followed by a description of the survey methodology used in the user study.

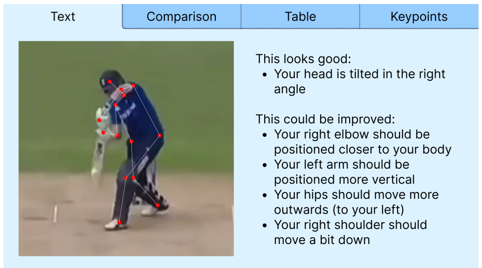
3.1 Prototypes

To investigate which explanation formats are most effective in a cricket-specific context, two prototypes were developed based on an XAI system that analyzes a cricket player's pose during a Cover Drive. An image from a cricket shot dataset was used as the user's pose and a frame from a YouTube tutorial video on performing a Cover Drives was used as the reference for an ideal pose [12].

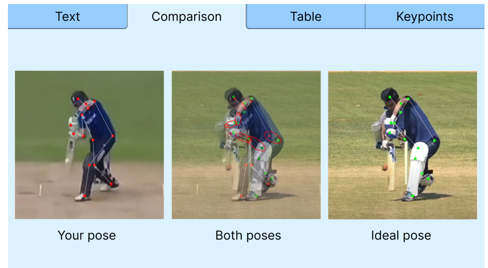
Following the recommendation to allow users to switch between different explanation formats [6], both prototypes were designed with tab-based navigation. This setup aimed to mimic an ideal learning environment in which users can compare and choose between explanation formats according to their personal preferences.

Four explanation formats were included, based on the categories proposed by Vilone and Longo [14]: textual, visual, rule-based, and mixed. The mixed explanation combined textual and visual elements as this pairing was considered most compatible and comprehensible for general users. Numeric explanations were excluded due to their lack of intuitiveness for general end-users. Each explanation was paired with an annotated input image, representing the output of a Human Pose Estimation model.

In the first prototype, the first explanation (*Text*) had a textual format, which included statements describing what the user did well and how they could improve their pose. The second explanation

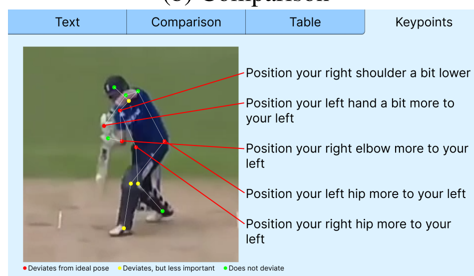


(a) Text



(b) Comparison

(c) Table



(d) Keypoints

Figure 1: The first prototype with (a) a textual explanation, (b) a visual explanation, (c) a rule-based explanation, and (d) a mix between a textual and visual explanation

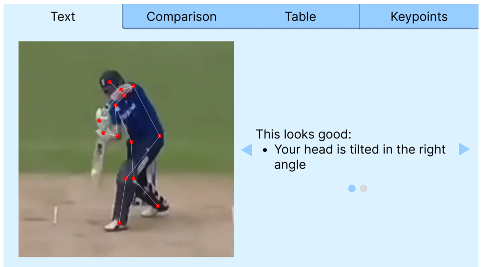
(*Comparison*) was of a visual format. The user's pose was presented next to the ideal pose, along with an image overlapping the two poses. In this overlay, key points that significantly deviated from the ideal were marked with red circles. The third explanation format (*Table*) had a rule-based format. The table in this explanation displayed positions or angles of important key points, ideal ranges in which these positions and angles should fall, and whether the user's data fell within these ideal ranges. The fourth and final explanation (*Keypoints*) integrated both textual and visual elements. Colored key points indicated feature deviation and importance, while a textual summary was given on how to improve the user's pose regarding key points that were both significant and incorrect. All explanations from the first prototype are illustrated in Figure 1.

A second prototype was developed to explore the impact of interactivity on explanation effectiveness. It used the same formats as the first prototype but added interactive elements. In the Text explanation, the text from the first prototype was split between two slides, one for positive feedback and one for improvements, to reduce cognitive load. In the Comparison explanation, additional annotated frames before and after the original frame were added to provide more context and pose information. In the Table explanation, hovering or clicking on rows in the table highlighted relevant key points to add more clarity. Finally, in the Keypoints explanation, users could click on key points in the image of their pose to reveal associated textual explanations one at a time. Examples from this second, interactive prototype are shown in Figure 2.

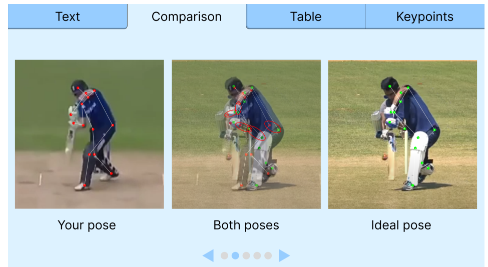
3.2 Survey

A user study was conducted using a digital survey to evaluate the two prototypes. Participants rated each explanation format on two dimensions: satisfaction and trust. These are two relevant measurements of the six measurements proposed by Hoffman et al. [7] to evaluate XAI models. To measure satisfaction, the Explanation Satisfaction Scale [7, Table 3] was used. Participants could respond to seven statements using a Likert scale between 1 and 5 (I agree strongly, I agree somewhat, I'm neutral about it, I disagree somewhat, I disagree strongly). The Explanation Satisfaction Scale was adapted to fit the cricket learning context, as shown in Table 1. To assess trust in the explanations, the XAI Trust scale [7, Table 8] was used. This scale works similarly to the Explanation Satisfaction Scale and was also adapted to a cricket-specific context, as shown in Table 2.

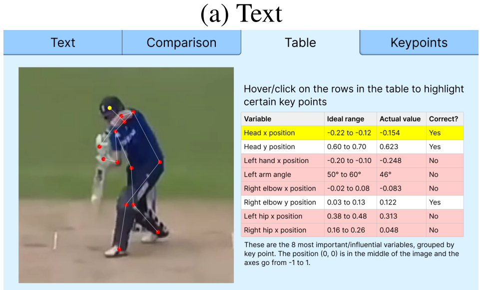
After evaluating each explanation within a prototype, users were asked which explanation(s) they would prefer to use in a real cricket learning environment. They were also invited to provide optional qualitative feedback and recommendations for each explanation and the overall prototype.



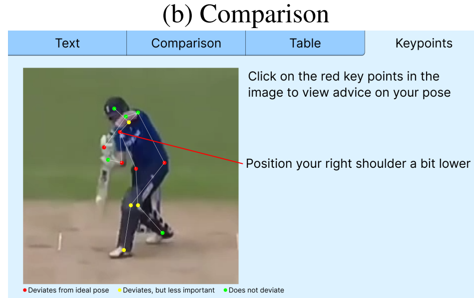
(a) Text



(b) Comparison



(c) Table



(d) Keypoints

Figure 2: Parts of the second, more interactive, prototype with (a) a textual explanation, (b) a visual explanation, (c) a rule-based explanation, and (d) a mix between a textual and visual explanation

Table 1: Adapted Explanation Satisfaction Scale

Q1	From this explanation, I know how to improve my pose.
Q2	This explanation of how to improve my pose is satisfying.
Q3	This explanation of how to improve my pose has sufficient detail.
Q4	This explanation of how to improve my pose seems complete.
Q5	This explanation tells me how I can use it to improve my pose.
Q6	This explanation of how to improve my pose is useful for learning cricket.
Q7	This explanation of how to improve my pose shows me how accurate the AI model is.

Table 2: Adapted XAI Trust Scale

Q1	I am confident in this explanation. I feel that it works well.
Q2	The outputs of this explanation are very predictable.
Q3	This explanation is very reliable. I can count on it to be correct all the time.
Q4	I feel safe that when I rely on this explanation I will get the right answers.
Q5	I am wary of this explanation.
Q6	This explanation is better than an explanation from a novice cricket player.
Q7	I like using this explanation for decision making.

Table 3: Average results for the first prototype

	Average	Q1	Q2	Q3	Q4	Q5	Q6	Q7
Text								
Satisfaction	3.71	4.25	3.83	3.42	3.50	4.33	4.08	2.58
Trust	3.37	4.00	3.33	3.17	3.33	2.83	2.92	3.67
Comparison								
Satisfaction	3.67	3.92	4.08	3.42	3.58	3.67	3.50	3.50
Trust	3.36	3.67	3.67	3.50	3.33	2.83	3.17	3.00
Table								
Satisfaction	2.54	1.75	1.67	3.25	3.00	2.58	1.75	3.75
Trust	2.54	2.42	2.83	3.08	2.92	3.67	2.00	2.17
Keypoints								
Satisfaction	3.95	4.42	4.25	3.83	3.75	4.42	3.92	3.08
Trust	3.45	4.00	3.50	3.08	3.58	3.08	3.42	3.67

Table 4: Average results for the second prototype

	Average	Q1	Q2	Q3	Q4	Q5	Q6	Q7
Text								
Satisfaction	3.62	4.25	3.58	3.33	3.25	4.42	3.92	2.58
Trust	3.35	3.83	3.50	3.25	3.33	3.00	3.08	3.42
Comparison								
Satisfaction	3.65	3.58	3.67	3.83	3.92	3.50	3.67	3.42
Trust	3.46	3.83	3.83	3.58	3.25	2.83	3.08	3.50
Table								
Satisfaction	2.86	2.25	2.17	3.50	3.33	2.33	2.33	4.08
Trust	2.79	2.58	3.17	3.42	3.25	3.33	2.17	2.25
Keypoints								
Satisfaction	3.96	4.33	4.33	3.83	3.75	4.33	4.08	3.08
Trust	3.73	4.25	3.67	3.33	3.67	2.50	3.50	4.17

4 Results

A total of twelve participants completed the survey. Their responses were evaluated using a 5-point Likert scale described in Subsection 3.2, where a score of 1 indicates a low and a score of 5 indicates a high satisfaction or trust. Notably, question 5 of the XAI Trust Scale was reverse-coded, as agreement with this negatively worded statement implies lower trust. For consistency in analysis, these responses were inverted when calculating average scores.

Descriptive statistics, consisting of the average scores per statement and the average ratings for each explanation, are presented in Table 3 for the first prototype without interactive explanation and in Table Table 4 for the second, more interactive, prototype. Participants were also asked to indicate which explanation(s) they would prefer to use in a real cricket learning environment. These preferences are summarized in Table 5.

To examine whether there were significant differences in user satisfaction and trust across all explanation formats, four one-way ANOVA tests were conducted for the satisfaction and trust scores for

Table 5: Amount of users that would prefer certain explanations in a cricket learning environment

	Prototype 1	Prototype 2
Text	5	5
Comparison	6	6
Table	0	1
Keypoints	11	11

both prototypes. For the first prototype, a significant difference was found in both satisfaction ($F(3) = 5.841$, $p = 0.00189$) and trust ($F(3) = 2.86$, $p = 0.0476$). For the second prototype, a significant difference was only found for the satisfaction scores ($F(3) = 5.429$, $p = 0.00289$) but not for the trust scores ($F(3) = 2.81$, $p = 0.0504$) using a significance level of 5%.

To identify which explanation formats contributed to these differences, Tukey's Honestly-Significant Difference post-hoc tests were performed. For the first prototype, the Table explanation scored significantly lower in satisfaction than all other explanations ($p < 0.05$). However, no significant differences were found in trust. For the second prototype, the Table explanation again scored significantly lower in satisfaction compared to the Comparison and Keypoints explanations ($p < 0.05$). However, the difference between the Table and Text explanation was not statistically significant.

To assess the impact of interactivity on the effectiveness of explanations, paired t-tests were conducted comparing the satisfaction and trust scores between the first and second prototype. No significant difference was found in satisfaction scores across the prototypes. However, the trust scores of the second prototype were significantly higher than the scores of the first prototype ($t(47) = 2.1621$, $p = 0.01787$).

5 Discussion

The results of this study suggest that the explanation format significantly influences the perceived effectiveness of an explanation. In particular, the Table explanation was consistently rated lower by participants compared to other explanations, as supported by the statistical analysis and the user preferences in Table 5. Qualitative feedback from participants further supports this finding: several users described the table as too complicated, which made it difficult for them to interpret the information and translate it into actionable feedback for improving their pose. These results imply that in a cricket-specific context, users may have a clear preference for explanation formats that are more intuitive, such as visual or textual formats, over rule-based formats.

In examining the role of interactivity, a significant increase in trust was found comparing the second to the first prototype. This finding aligns with previous research highlighting the role of interactive elements on cognitive load [3]. However, it contrasts with findings from other domains, such as recommender systems, where interactivity did not translate to more effective explanations [10]. This discrepancy suggests that the effects of interactivity may be highly context-dependent.

Interestingly, while a significant difference in trust was found between the two prototypes, no such difference was found in satisfaction. This indicates a potential trade-off between transparency and usability. For example, one user gave feedback that they found the interactive Keypoints explanation quite intuitive but disrupted by the amount of required clicking.

It is important to acknowledge the limitations of this study. The sample size was relatively small ($N = 12$). This may affect the generalizability of the findings. Future studies with larger and more diverse participant groups are necessary to validate the results of this study and draw more definitive conclusions.

Furthermore, all explanations used in the prototype were manually created and may not fully reflect the actual outputs of an XAI model. Providing more accurate and model-generated explanations could yield different results and offer more accurate insights into how users perceive explanations.

In summary, this study demonstrates that explanation formats could significantly shape a user's perception of given explanations in a cricket learning context. Moreover, interactive features could possibly increase user trust, though this does not necessarily equate to higher explanation satisfaction.

6 Responsible research

The survey data for this study was collected from a small group of participants who verbally consented to the storage and analysis of their responses. The collected data was limited to functional and task-relevant input, with no personally identifiable data gathered. Nevertheless, no formal ethical procedures or institutional review processes were followed during data collection. For future studies, particularly those involving a larger group of participants, it is recommended that such ethical protocols are implemented to safeguard the privacy of the participants and ensure research integrity.

Another ethical consideration relates to the possible bias inherent in AI systems that classify and analyze cricket shots. While the goal of XAI models is to increase transparency in otherwise opaque AI models, bias could still occur. Although the prototypes used in this study did not utilize actual (X)AI models, they are intended to simulate the output of such models and may be repurposed in future applications involving real (X)AI models. It is therefore essential to communicate clearly to users that outputs shown in such prototypes may be inaccurate or biased.

To support reproducibility and facilitate future research, this paper includes visual documentation of the prototypes used, as well as the instruments used for the survey. Aggregate quantitative results are also reported to allow for comparison in follow-up studies. However, individual-level data has not been disclosed to protect participant anonymity.

7 Conclusion and future work

This study explored the effectiveness of different explanation formats in the context of cricket training. Two prototypes were developed, each featuring four different explanation formats (textual, visual, rule-based, and mixed) on how to improve the pose of a player whilst performing a cricket shot. Numeric explanations were excluded from the prototypes due to their lack of intuitiveness. The second prototype introduced interactive features in its explanations, while the first one relied on static presentations.

Findings from a small-scale user study suggest that users expressed a clear preference for textual, visual, and mixed explanations over the rule-based explanation, which was rated significantly lower in terms of satisfaction. However, no significant individual differences were observed regarding trust in the explanations.

When comparing the two prototypes, results indicate that the inclusion of interactive features led to a statistically significant increase in user trust, though this did not translate to significantly higher satisfaction.

This study could be extended in a few different ways. One key extension would be the integration of an actual (X)AI pipeline, thereby allowing for more accurate explanations and details. Furthermore, conducting the study with a larger participant group would improve the reliability and generalizability of the findings. Further research could also explore the effects of different types of interactive features on the effectiveness of explanations.

Another way to extend this research is by expanding the scope of the system beyond Human Pose Estimation. By, for example, also estimating the position, direction, and speed of the ball or estimating the speed and direction of the cricket bat, an XAI model could give more accurate explanations and comprehensive feedback.

References

[1] Waqas Ahmad, Muhammad Munsif, Habib Ullah, Mohib Ullah, Alhanouf Abdulrahman Alsuwailem, Abdul Khader Jilani Saudagar, Khan Muhammad, and Muhammad Sajjad. Optimized deep learning-based cricket activity focused network and medium scale benchmark. *Alexandria Engineering Journal*, 73:771–779, July 2023. ISSN 1110-0168. doi: 10.1016/j.aej.2023.04.062. URL `https://www.sciencedirect.com/science/article/pii/S1110016823003368`.

[2] Housam Khalifa Bashier Babiker and Randy Goebel. An Introduction to Deep Visual Explanation, March 2018. URL `http://arxiv.org/abs/1711.09482`. arXiv:1711.09482 [stat].

[3] Cintia Bali, Gergő Várkonyi, Mónika Szabó, and András N. Zsidó. The impact of visual cues on reducing cognitive load in interactive storybooks for children. *Journal of Experimental Child Psychology*, 260: 106320, December 2025. ISSN 00220965. doi: 10.1016/j.jecp.2025.106320. URL `https://linkinghub.elsevier.com/retrieve/pii/S0022096525001262`.

[4] Tanmay Chakraborty, Marion Koelle, Jörg Schlötterer, Nadine Schlicker, Christian Wirth, and Christin Seifert. Explanation format does not matter; but explanations do – An Eggsbert study on explaining Bayesian Optimisation tasks, April 2025. URL `http://arxiv.org/abs/2504.20567`. arXiv:2504.20567 [cs].

[5] Ks Sreekar Datta, G Narasimha Naidu, Rahul Dasari, and T V Jayakumar. Advancements in Cricket Shot Detection: Integrating Human Pose Estimation and Deep Learning for Automated Analysis. In *2024 15th*

International Conference on Computing Communication and Networking Technologies (ICCCNT), pages 1–7, June 2024. doi: 10.1109/ICCCNT61001.2024.10726249. URL https://ieeexplore.ieee.org/document/10726249. ISSN: 2473-7674.

[6] Mengisti Berihu Girmay and Felix Möhrle. Exploring explainability formats to aid decision-making in dairy farming systems. pages 269–274. Gesellschaft für Informatik e.V., 2024. ISBN 978-3-88579-738-8. URL https://dl.gi.de/handle/20.500.12116/43885. ISSN: 2944-7682.

[7] Robert R. Hoffman, Shane T. Mueller, Gary Klein, and Jordan Litman. Measures for explainable AI: Explanation goodness, user satisfaction, mental models, curiosity, trust, and human-AI performance. *Frontiers in Computer Science*, 5, February 2023. ISSN 2624-9898. doi: 10.3389/fcomp.2023.1096257. URL https://www.frontiersin.org/journals/computer-science/articles/10.3389/fcomp.2023.1096257/full. Publisher: Frontiers.

[8] Pigi Kouki, James Schaffer, Jay Pujara, John O'Donovan, and Lise Getoor. User Preferences for Hybrid Explanations. In *Proceedings of the Eleventh ACM Conference on Recommender Systems*, RecSys '17, pages 84–88, New York, NY, USA, August 2017. Association for Computing Machinery. ISBN 978-1-4503-4652-8. doi: 10.1145/3109859.3109915. URL https://dl.acm.org/doi/10.1145/3109859.3109915.

[9] Pigi Kouki, James Schaffer, Jay Pujara, John O'Donovan, and Lise Getoor. Personalized explanations for hybrid recommender systems. In *Proceedings of the 24th International Conference on Intelligent User Interfaces*, IUI '19, pages 379–390, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 978-1-4503-6272-6. doi: 10.1145/3301275.3302306. URL https://dl.acm.org/doi/10.1145/3301275.3302306.

[10] Pigi Kouki, James Schaffer, Jay Pujara, John O'Donovan, and Lise Getoor. Generating and Understanding Personalized Explanations in Hybrid Recommender Systems. *ACM Trans. Interact. Intell. Syst.*, 10(4): 31:1–31:40, November 2020. ISSN 2160-6455. doi: 10.1145/3365843. URL https://dl.acm.org/doi/10.1145/3365843.

[11] Hafeez Ur Rehman Siddiqui, Faizan Younas, Furqan Rustam, Emmanuel Soriano Flores, Julién Brito Ballester, Isabel de la Torre Diez, Sandra Dudley, and Imran Ashraf. Enhancing Cricket Performance Analysis with Human Pose Estimation and Machine Learning. *Sensors*, 23(15):6839, January 2023. ISSN 1424-8220. doi: 10.3390/s23156839. URL https://www.mdpi.com/1424-8220/23/15/6839. Number: 15 Publisher: Multidisciplinary Digital Publishing Institute.

[12] SIKANA English. How to Play a Cover Drive | Cricket, May 2017. URL https://www.youtube.com/watch?v=fELk4k1pndg.

[13] Kevin Tissera, Dominic Orth, Minh Huynh, and Amanda C. Benson. The impact of augmented feedback (and technology) on learning and teaching cricket skill: A systematic review with meta-analysis. *PLOS ONE*, 17(12):e0279121, December 2022. ISSN 1932-6203. doi: 10.1371/journal.pone.0279121. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0279121. Publisher: Public Library of Science.

[14] Giulia Vilone and Luca Longo. Classification of Explainable Artificial Intelligence Methods through Their Output Formats. *Machine Learning and Knowledge Extraction*, 3(3):615–661, September 2021. ISSN 2504-4990. doi: 10.3390/make3030032. URL https://www.mdpi.com/2504-4990/3/3/32. Number: 3 Publisher: Multidisciplinary Digital Publishing Institute.

[15] World Population Review. Cricket Countries 2025. URL https://worldpopulationreview.com/country-rankings/cricket-countries.

[16] Weronika Łajewska, Damiano Spina, Johanne Trippas, and Krisztian Balog. Explainability for Transparent Conversational Information-Seeking. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '24, pages 1040–1050, New York, NY, USA, July 2024. Association for Computing Machinery. ISBN 979-8-4007-0431-4. doi: 10.1145/3626772.3657768. URL https://dl.acm.org/doi/10.1145/3626772.3657768.