www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 2.1, (2026), pp. 7161-7175
Open Access. Online & Print

DOI: 10.5281/zenodo.19671118

ARTIFICIAL INTELIGENC SYSTEMS FOR AMATEUR CYCLIST

José Párraga-Valle[1], Jaime Meza[1], Gustavo Chafla-Altamirano[2*], Nelson Salgado-Reyes[2], Raisa Socorro-Llanes[3] and Natali Elizabeth Lascano Robalino[4]

[1]*Departamento de Sistemas Computacionales, Universidad Técnica de Manabí (UTM), Portoviejo, 130105, Ecuador*
jose.parraga@utm.edu.ec
https://orcid.org/0000-0001-6414-2268
jaime.meza@utm.edu.ec
https://orcid.org/0000-0003-4960-7957
[2]*Facultad de Hábitat, Infraestructura y Creatividad, Carrera Sistemas de Información, Pontificia Universidad Católica del Ecuador (PUCE), Quito, 170129, Ecuador*
gxchafla@puce.edu.ec
https://orcid.org/0000-0003-4754-4446
nesalgado@puce.edu.ec
https://orcid.org/0000-0001-8908-7613
[3]*Universidad Tecnológica de La Habana José Antonio Echeverría (CUJAE), Cuba*
raisa@ceis.cujae.edu.cu
https://orcid.org/0000-0002-2627-1912
[4]*Facultad de Ingeniería y Ciencias Aplicadas, Carrera de Ingeniería en Sistemas de Información, Universidad Central del Ecuador (UCE), Quito, 170521, Ecuador*
nelascano@uce.edu.ec
https://orcid.org/0009-0009-3832-2965

Received: 04/01/2026
Accepted: 28/03/2026

Corresponding Author: Gustavo Chafla Altamirano (gxchafla@puce.edu.ec)

ABSTRACT

Recommender systems (RS) have yielded significant results in various fields of knowledge and have great potential in the context of sports activities. In the context of amateur cycling, the use of machine learning has enormous potential to improve recommender systems applied to solving problems in amateur cycling. However, there is still no clear understanding of how artificial intelligence techniques have been used, due to the lack of secondary studies in this context. Therefore, we conducted a Systematic Literature Review (SLR) with the goal of understanding how artificial intelligence has been used in recommender systems and investigating the real evidence of benefits and/or difficulties encountered in their use. Nowadays, approximately 72% of the studies presented positive evidence of benefits. However, approximately 59% presented positive evidence of difficulties. These findings indicate that the potential of RS in amateur cyclists needs to be explored. Therefore, we conclude that the results obtained in this systematic literature review are very important for the computer science and sports communities, as they present an overview of the main

studies published on how RS is used by amateur cyclists. In future work, we intend to expand this SLR to include new contexts and more specific research questions.

KEYWORDS: Systematic Literature Review, Machine Learning, Recommender Systems, Cycling, Amateur, Sport.

1. INTRODUCTION

The use of recommender systems (RS) has gained significant relevance in multiple areas, from e-commerce to digital content consumption, due to their ability to personalize experiences and facilitate decision-making (Yang & Zhao, 2025). In sports, and particularly in recreational or amateur cycling, their application is still incipient, despite the significant growth of this practice worldwide.

Various digital platforms and wearable devices have enabled the collection of large volumes of data related to routes, training plans, health, physiological metrics, anthropological data of users, and environmental conditions (Wen et al., 2025). However, most current systems lack specific models that address the needs of amateur cyclists, who present heterogeneous profiles in terms of experience level, motivation, nutrition, personal goals, and physical limitations. This gap highlights the lack of solutions that integrate personalized and accessible approaches for this user group (Breyer et al., 2025).

For this reason, a Systematic Literature Review (SLR) (García-Peñalvo, 2022; Lee et al., 2023) was conducted to investigate and understand how SR has been used with amateur cyclists. This review provides a new perspective on the use of systems by applying machine learning techniques through the analysis of the main studies presented in the literature. The main research question is: How has SR been used with amateur cyclists? As a secondary objective, the SLR seeks to investigate the real evidence of the benefits and/or difficulties encountered in the process of using SR with amateur cyclists. This article is composed of five sections. Section I presents an introduction to the proposed work. Next, Section II illustrates the developed SLR protocol. Section III presents the execution process. Section IV shows the results obtained and presents some discussions. Finally, Section V presents some conclusions about this study.

2. METHOD

In developing this study, we used the strategy of designing and conducting a Systematic Literature Review (SLR), which is a form of secondary research that uses a well-defined methodology to identify, analyze, and interpret all available evidence on a specific research question in an unbiased and repeatable manner (Thyago et al., 2021). An optimal systematic literature review requires time and effort, which must follow the established methodology to avoid bias.

It should be noted that, in a Systematic Review, researchers search for and analyze studies published in the literature, indicating both those related to their research questions and those unrelated, excluding the latter from the process. In this sense, our systematic literature review followed the guidelines and protocol for SLR presented by Thyago et al. (2021). We also used the template used in a previous review conducted by the authors. Using the guide and templates allows for greater agility in the review process. The guidelines indicate some steps to follow to conduct a successful SLR. The first step is to create a document called a Protocol. This file contains all the information necessary to guide the execution process, including the objectives, research questions, keywords and synonyms, search string, databases, selection criteria (inclusion, exclusion, and quality), the extraction form, and the search process, among other information (Thyago et al., 2021). This section briefly presents the protocol developed for this SLR.

The main objective of the RLS is to understand how the concept of Recommender Systems (RS) has been used by amateur cyclists, providing an overview of the use of RS applied to amateur cycling. This research seeks to answer the following research question: How have recommender systems been used by amateur cyclists? To further specify the research question, our main focus is to analyze the whys and wherefores of RS use by amateur cyclists, as well as to seek evidence of the benefits it provides. Furthermore, this review will also address the difficulties encountered in the process of using RS. Finding common patterns and metrics of RS use in cycling can help us map the various fields in which this concept can be applied.

After defining the objectives of the RLS, the second step is to create the research questions. These were created based on the predefined objectives and are presented in Table I. For each question, we describe the motivation for creating it and what it aims to answer. It should be noted that RQ1 seeks to answer the main objective of the SLR, while RQ2 and RQ3 seek to answer the secondary objectives. The third step is to define the search strategy. I this case, the first, the main keywords (and their synonyms) are defined to generate the search string. In this sense, the keywords (and their synonyms) for this SLR are: G1 – RS Context: Recommendation System, Recommendation Algorithms, G2 – RS Metrics: Deep Learning, Prediction Techniques, Prediction Model.

We can see that the keywords are divided into three groups (G1, G2, and G3). Group G1 refers to words related to the term "Recommend System" and its possible synonyms. Group G2 refers to words related to Machine Learning and its synonyms. Finally, G3 refers to words related to "Cyclist" and its synonyms. It should be noted that, in this context, "synonym" refers to terms related to the main keyword. Since we need any combination of keywords G1 with keywords G2 and G3, we will use the OR operator between words from the same group and the AND operator between words from different groups. Therefore, we obtain the following search string construction:

1. "Recommender System"
2. "Recommendation System*" OR "Recommendation Algorithms"
3. "Machine Learning" OR "Deep Learning"
4. "Prediction Techniques" OR "Prediction Model"
5. "Cyclist*" OR "Cycling"
6. "Sport*" OR "Athlete*"

The summarized search string is "(1 OR 2) AND (3 OR 4) AND (5 OR 6)". The complete and defined base search string for this systematic review is presented below. It is important to note that minor variations were made to the search string to obtain the correct studies depending on the database used, as they use different search engines, and the final search string = ("Recommender System*" OR "Recommendation System*" OR "Recommendation Algorithms") AND ("Machine Learning" OR "Deep Learning" OR "Prediction Techniques" OR "Prediction Model") AND ("Cyclist*" OR "Cycling" OR "Sport*" OR "Athlete*"). The asterisk (*) is also included to include plural terms.

The third step was defining the databases (DBs) in which we will apply the search string to obtain the primary studies. We will consider databases in the computing field that have article availability, allow keyword searching, and have high bibliographic relevance. In this regard, we have selected the following databases: SCOPUS, Web of Science (WOS), and IEEE Digital Library.

The fourth step was the selection criteria (inclusion, exclusion, and quality), which will serve as the basis for the entire study selection process. These criteria are based on all the previously defined information and aim to improve the results obtained through SLR (Mendez et al., 2024). These criteria can be inclusion criteria (minimum criteria that articles must meet to be included), exclusion criteria (criteria that eliminate a study from the selection), and quality criteria (criteria that classify the selected studies according to their quality).

The selection criteria (inclusion and exclusion) defined for this systematic review were based on those defined in the study (Thyago et al., 2021), with minor modifications. In summary, the inclusion criteria are primary and peer-reviewed studies that use the Recommender System applied to sport and meet a minimum quality of 50%. It is important to note that a study must meet all inclusion criteria to be selected for SLR. However, one exclusion criterion is sufficient to remove it from the SLR. The main exclusion criteria for this SLR include non-primary, duplicate, incomplete, domain-specific, or non-peer-reviewed studies, short articles, languages other than English, and out of scope (Thyago et al., 2021). Regarding the quality criteria, we reused those presented in the study (Thyago et al., 2021). The quality criteria defined for this SLR consist of a set of 12 questions that evaluate aspects of the analyzed studies, such as reasoning, clarity of objectives, proposed models, techniques, and metrics, presentation of results, and limitations. The description of the application context and the possibility of expansion to other contexts, as well as the existence of tools and/or the evaluation of proposals, are also evaluated. Each question is assigned a score. We adapt the possible answers to maintain the same scale across all questions (Yes = 1.0, Partially = 0.5, and No = 0.0). Thus, an article can have a quality score ranging from 0 to 12 (maximum quality). The cutoff point was set at 6 (50%), inclusive.

The fifth step is to define the article selection process. The process initially consists of running the search string in the selected databases. The results are exported to compatible files and imported into the ZOTERO bibliographic management tool.

According to Thyago et al. (2021), after adding the studies to the tool, the following steps should be followed:

1. Remove duplicate studies.
2. Read the titles, abstracts, and keywords of the studies, excluding those that meet the exclusion criteria.
3. Read the introduction and conclusion of the studies approved in step 2, excluding those that meet the exclusion criteria.
4. Obtain the full version of the selected primary studies (usually in PDF format) from the databases.
5. Analyze the quality of the studies according to the quality criteria, excluding those with a quality score below 50%.
6. Read the full study and extract the necessary data according to the extraction form.

It is important to note that any questions about the inclusion of an article in a specific step will be approved for the next step. At the end of the process, the remaining articles will be submitted for data extraction. Step 5 of the process consists of reading all the studies selected in the previous step. This step involves extracting the data necessary for the general and specific analyses that meet the objectives of the SLR. To do this, an extraction form is created containing all the data to be extracted from the selected studies (Thyago et al., 2021). Table II shows that the first five pieces of data refer to general information about the studies, while the last four refer to data to answer our research questions. Both are important for the results analysis process, described in Section IV.

Table 1: Research Questions of the Systematic Literature Review.

Research Question (RQ)	Description and Motivation
RQ-01: What are the main reasons for using the Recommender System with amateur cyclists?	The objective is to understand the problems addressed in the study, seeking to answer what problems SR helped solve and how it was used in amateur cycling.
RQ-02: What are the benefits of using the Recommender System with amateur cyclists?	This question focuses on whether the studies analyzed provide evidence of the benefits generated by the use of SR in amateur cycling.
RQ-03: What are the challenges faced by using the Recommender System with amateur cyclists?	This question focuses on whether the studies analyzed provide evidence of the difficulties generated by the use of SR in amateur cycling.

3. SYSTEMATIC LITERATURE REVIEW EXECUTION.

This began with running the search string in each selected database. To do this, we conducted an initial study of each search engine's performance. The same search string was applied to all three databases, given their specific requirements. Fortunately, all databases allowed exporting results—for SCOPUS and WOS—to RIS and IEEE Digital Library files in CSV format, which facilitated the extraction and import of results into the ZOTERO tool. ZOTERO is a personal research assistant, a free and easy-to-use tool that helps you collect, organize, annotate, cite, and share research to build an SLR. As stated on its website, "In addition to providing a way to document the entire process, the tool will help you remember what's important during a systematic literature review" (Zientarska-Kayko et al., 2024).

According to Table III, the database with the most articles was Scopus (90), followed by ISI Web of Science with 74. Finally, the last database was the IEEE Digital Library with only 17 studies. In total, 181 studies were selected and extracted. Figure 1 shows the same data in percentages for better visualization.

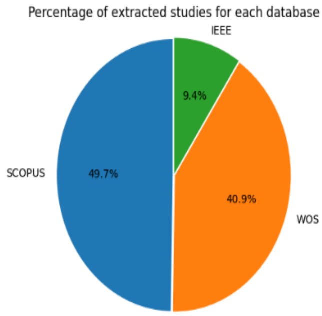

Fig 1: Percentage of studies extracted for each database.

The majority came from Scopus (49.6%), followed by Web of Science (40.9%). The database with the fewest returned studies was the IEEE Digital Library, with only 9.4% of the studies. The 181 studies were retrieved and loaded into the ZOTERO tool using RIS files (SCOPUS and WOS) and CSV files for the IEEE database. The process defined in Section II was then executed. Figure 2 presents a summary of the process and the results obtained in each step. Figure 2 shows that the first step was to obtain the studies from the databases and group them in the ZOTERO tool to optimize process management. Using the ZOTERO tool, duplicate studies were eliminated (a total of 50 studies), leaving a total of 131 studies for analysis.

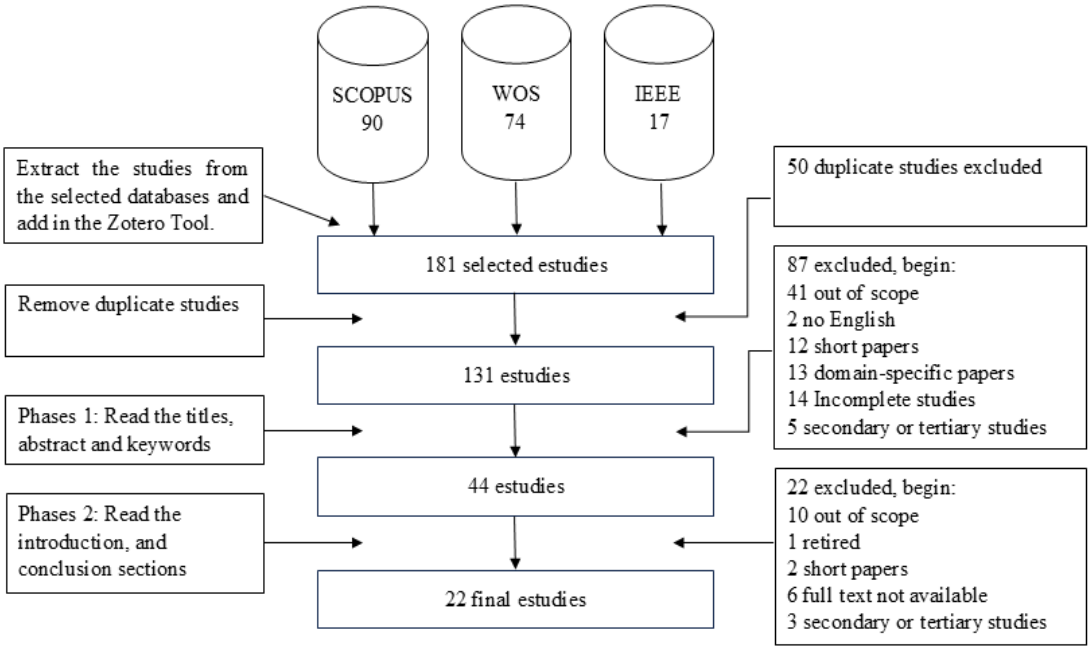

Fig. 2: Summary of the performed process. Based on (Mendez et al., 2024).

From this point, we start with the documents screening; in this sense, the first one consisted of reading the titles, abstracts, and keywords of the studies. After the first stage, 87 studies were rejected for the following reasons: 41 studies that did not use SR in amateur cyclists (out of scope), two articles written in a language other than English, 12 short articles (less than four pages), 13 domain-specific articles, 14 incomplete studies, and five secondary or tertiary studies. At the end of Phase 1, 44 studies remained.

Table: 2

Number	Study Information	Description and possible values
1	Study ID	Number used to identify the study (index - (template Sxx))
2	Authors, year, title, and country	General information about the study
3	Type of Study	A journal, workshop, conference, or other.
4	Context of Study	Industry, academia, or both
5	Research Method (Mendez et al., 2024; Shull et al., 2008)	Controlled experiment, Case study, Survey, Ethnography, Action research, Illustrative scenario, not applicable
6		Why and for what purpose should RS be used by amateur cyclists?
7	PI1 Main Objectives	Evidence (positive or negative) of the benefits provided by the use of RS, with or without empirical evaluation
8	PI2 Evidence of Benefits	Evidence (positive or negative) of difficulties encountered in the process of using RS, with or without empirical evaluation.

Phase 2 then begins, which consists of reading the introduction and conclusion of the studies. After this phase, 22 studies were rejected for the following reasons: 10 studies that did not use SR in amateur cyclists (out of scope), 1 study withdrawn due to journal policies, two short articles (<4 pages), six full texts unavailable online, and three secondary or tertiary studies.

Finally, the studies underwent a quality assessment. The 22 resulting studies are shown in Table III with their respective quality assessments (numbers are rounded). Once the studies have been correctly selected and classified, the final step of the selection process begins, which consists of reading the full text of the selected studies and extracting the general and specific data indicated in the extraction form (Table II). A data analysis was performed, and the results are illustrated in Section IV.

Table 3: STUDIES SELECTED AND THEIR QUALITY ASSESSMENT.

TABLE III					
STUDIES SELECTED AND THEIR QUALITY ASSESSMENT					
Id	**Author**	**%**	**Id**	**Author**	**%**
S01	(Ramaraj et al., 2024)	80%	S12	(GUO, 2025)	70%
S02	(Rostami et al., 2024)	60%	S13	(Hassan et al., 2024)	55%
S03	(Deepak et al., 2023)	70%	S14	(Zhao, 2024)	60%
S04	(Meng & Qiao, 2023)	50%	S15	(X. Wang et al., 2023)	70%
S05	(Chikov et al., 2022)	55%	S16	(Mahyari & Pirolli, 2021)	70%
S06	(Zhang, 2025)	85%	S17	(Deepak & Anguraj, 2023)	65%
S07	(Andric et al., 2021)	65%	S18	(Mahyari et al., 2022)	75%
S08	(Kumar & Sahani, 2023)	65%	S19	(Deepak et al., 2022)	90%
S09	(Demosthenous et al., 2022)	90%	S20	(Lijuan et al., 2025)	90%
S10	(T. Wang & Park, 2021)	60%	S21	(Liu & Cao, 2024)	65%
S11	(Chen & Tian, 2024)	60%	S22	(Feely et al., 2023)	70%

4. DATA ANALYSIS DISCUSSION

A complete analysis of the results of the 22 studies was conducted. Data were extracted from the data extraction form developed in Microsoft Excel, as shown in Table II, using the ZOTERO tool. This tool allows for annotation and highlighting of the most important ideas for better analysis and manipulation. First, a descriptive analysis of the data was performed. This analysis is presented in Subsection IV-A. Subsequently, a more detailed analysis was performed to answer the research questions. The latter is illustrated in Subsection IV-B.

4.1. Descriptive data analysis

The first step was to analyze the studies in general terms using descriptive analysis. This type of analysis seeks to identify useful information in the data and, as a result, provides summaries of the study samples. The analysis of the studies involved the following variables: study quality, year of publication, study source, authors' country, context of application, research method used, and educational level. The results and discussion are described in the following paragraph.

TABLE: 5		
QUALITY ASSESSMENT RESULTS		
Qual. (%)	**Quantity**	**Papers**
50%	1	S4
55%	2	S5, S13
60%	4	S2, S10, S11, S14
65%	4	S7, S8, S17, S21
70%	5	S3, S12, S15, S16, S22
75%	1	S18
80%	1	S1
85%	1	S6
90%	3	S9, S19, S20

Variable of Study Quality: The first analysis was conducted using the study quality metric. It should be noted that this metric corresponds to a percentage (from 0 to 100%) that indicates the score obtained by the study, based on the score assigned to each quality question defined in this SLR. A score of 0% indicates low-quality studies, and a score of 100% indicates the highest possible quality study. Figure 3 illustrates the number of studies according to the score assigned. Figure 3 shows variation in quality scores among the selected studies, with emphasis on scores of 60, 65, and 70 (4, 4, and 5 studies, respectively). The highest score was 90%, obtained by three studies. Overall, the average score awarded was 69.09, and the median was 67.5. Table IV provides more details on the studies and their quality scores.

The data shown in Table IV indicate that three studies (S04, S05, S13) obtained a quality assessment considered low (<60%). Followed by 13 studies (S02, S10, S11, S14, S07, S8, S17, S21, S3, S12, S15, S16, S22) obtained an average quality assessment (<75%). Another 3 articles (S18, S1, S6) presented good quality (<85%). Finally, only three articles (S9, S19, S20) presented a high quality (>=85%). Overall, most of the articles presented (16) had a quality considered low-medium (score <75%). This shows us that it is important for studies published in the literature to clearly present the reasoning, objectives, proposed techniques, results, limitations, context of application, and the possibility of expansion to other contexts, as well as the existence of tools and/or evaluation of proposals.

General variables (Year, source, and country): Year of publication. During the search process, it was limited to the last 5 years of publication, that is, from 2021 to 2025. Figure 4 presents the year of publication of the studies. The figure shows a smaller number of studies from 2021 (2 or 9.1%), for 2022, it shows a slight growth trend of studies (5 or 22.7%), with 2023 being the year with the most studies (7 or 31.8%). It should be noted that the data were obtained in mid-2025. The number of studies for 2024 equaled that of 2022 (5 or 22.7%). Finally, the year 2025 presents studies with (3 or 13.6%).

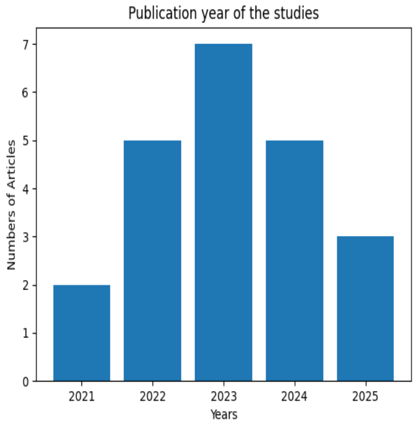

Fig. 3. Publication year of the studies.

The ***type of study source***. Figure 4 shows the sources of the studies. It is observed that the majority of the studies came from scientific articles (around 68.2%). Secondly, 31.8% of the studies came from conferences or congresses.

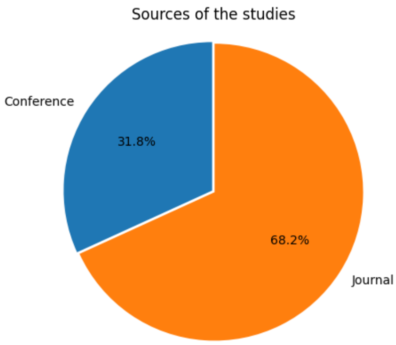

Fig. 4: Sources of the studies.

The country of the authors of the studies, always considering the first author. Table V shows the data by country. It is observed that the majority of the studies were published by authors from universities in China (8 or 36.4%) and India (5 or 22.7%), followed by authors from the USA (2 or 9.1%). Finally, the countries with only one study were Cyprus, Finland, Ireland, Italy, Pakistan, Russia, and South Korea (1 or 4.5% each). When analyzing the figures by year of publication of the studies, we observe a slight upward trend in the use of RS among amateur cyclists, but it is not possible to affirm that a trend exists in this regard. However, it can be observed that the majority of publications are scientific articles, which indicates recent research in the field. While China is the country with the most publications, other countries such as India and the USA also have significant publications. It should be noted that in this SLR, we only analyzed studies published in English, which creates a bias based on the language of publication.

TABLE 5: Studies By Country.

TABLE V STUDIES BY COUNTRY	
Country	***Number***
China	8
India	5
USA	2
Chipre, Finlandia, Ireland, Italy, Pakistan, Rusia, South Korea	1

Detailed variables: Application context, research method, and educational level:

The research method was analyzed and the context of application (as shown in the extraction form). In addition, we attempted to identify the educational level used in the research context. However, many authors do not clearly state this information in the text. The fifth variable analyzed was the context of application (academic, industrial, or both). The context of application is academic if the research was carried out in a school or university, usually with students and faculty from the institution. On the other hand, the research is considered industrial if it was carried out in the context of companies or businesses. Finally, the context can be both if the research is applied to the two previous cases. Figure 5 illustrates the data obtained in percentage terms.

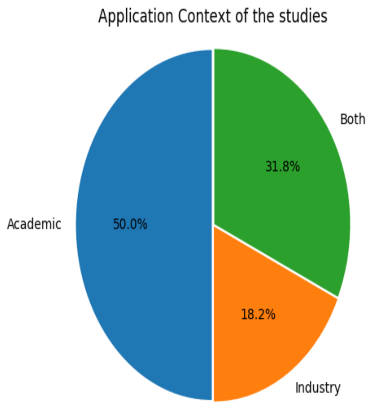

Fig. 6. Application Context of the studies.

When analyzing the context of application of the studies, it is observed that most are conducted in the academic field. In fact, given the more recent research in this area, the initial validations and results come from the academic field and are subsequently transferred to the industrial field. However, it is important to emphasize the importance of clearly describing contextual information in the study, as a significant number of studies did not present it.

The research method used, according to Mendez et al. (2024). A research method is a set of organizing principles around which empirical data are collected and analyzed (Mendez et al., 2024). There are a large number of research methods that can be applied, including controlled experiments, case studies, and surveys. It is also possible to combine several methods to form a mixed research method (Figure 7).

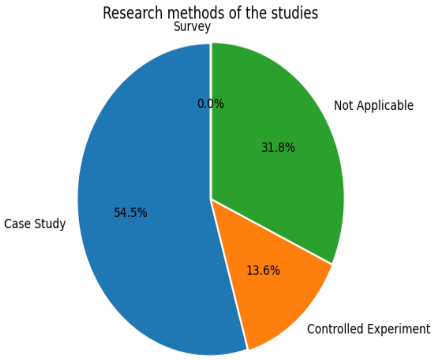

Fig. 7. Research methods of the studies.

Among the research methods, case studies were the most widely used. However, it is worth noting that a third of the studies did not use research validation methods, which represents a negative point for the community, as such evidence is essential to ensure the proper functioning of the developed approaches. It should be noted that this SLR study did not include other SLR studies focused on the research topic.

Another variable studied was educational level. This could be elementary, primary, secondary, high school, or university. Unfortunately, a third of the studies did not clarify the educational level used. Figure 8 illustrates the data obtained in percentage terms.

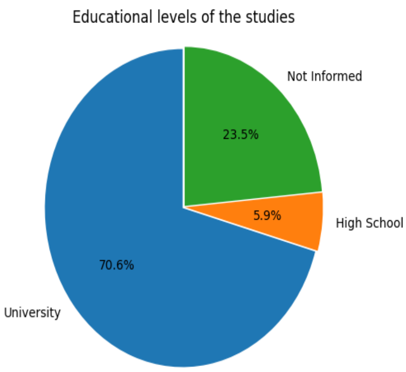

Fig. 8. Educational levels of the studies

It is observed that 17 (89.5%) of the studies worked at the university (S2, S3, S5, S6, S7, S8, S9, S10, S11, S12, S13, S14, S15, S16, S17, S18, S19, S20, S21, S22). On the other hand, only 1 (5.3%) (S1) worked, in addition to the university level, in secondary school. Finally, only 1 (5.3%) (S4) did not present information on their educational level. It should be noted that in most of the studies presented, the authors work in a higher education center, which is important to better understand the research developed using RS in amateur cyclists.

4.2. Research questions responses

Regarding Research Question 1 (RQ1). What are the main reasons for using SR with amateur cyclists? Its objective is to understand the problem addressed in the study, seeking to answer what problems SR helped solve and how it was used with amateur cyclists. It is clear that SR can be used to solve different problems in a variety of ways. With this question, we sought to analyze whether there were groups of problems with similar characteristics, in order to discover patterns of SR use among amateur cyclists. Two groups were identified: G1 (Context); G2 (Metrics)

The first group, G1, refers to the Context of the Recommender System. This group includes studies that, in some way, utilized the power of Machine Learning to improve SR among amateur cyclists. For a better understanding of the studies, group G1 is subdivided into 5 subgroups: G1-Subgroup 1-Training plan, G1-Subgroup 2-Nutrition, G1-Subgroup 3-Motivation, G1-Subgroup 4-Injuries, G1-Subgroup 5-Education.

G1-Subgroup 1-Training plan.- Study S3 proposes a model that offers a better balance compared to conventional approaches such as SVM, RF, k-NN, and logistic regression to automatically predict an individual's condition with characteristics such as age, sex, calories, temperature, blood pressure, heart rate, pulse, blood sugar level, respiratory conditions, and physical fitness (Deepak et al., 2023). Study S5 presented a recommendation system for measuring quantitative physiological parameters at the anaerobic threshold, and a predictive parameter model for a fatigue stress test using artificial intelligence (Chikov et al., 2022). Study S6 not only creates a system to help coaches and athletes understand physical fitness in real time, but also provides personalized and scientifically sound training recommendations based on training goals and actual needs, which is critical for improving athletic performance and preventing sports injuries (Zhang, 2025). In the same line, the Study S7 presented a recommendation system, using a matrix factorization approach with a personalized normalization solution to improve sport climbing, route exploration, and route search. The difficulty of a route, or its grade, is typically assessed by expert climbers, called routemakers. A regular climber, after trying a route may perceive it as more or less difficult than the route maker. It is important to estimate the climber's perceived difficulty of the routes in order to suggest routes with the target difficulty expected by the cyclist (Andric et al., 2021). Study S9 proposes a model capable of predicting a cyclist's speed, as well as other relevant target variables, with high accuracy. Another important contribution of this project is the creation of a recommendation module that provides cyclists with real-time suggestions to increase their average speed during the ride, with minimal impact on their average heart rate (Demosthenous et al., 2022). Study S10 presented an intelligent sports management system based on deep learning technology, utilizing information technology and human-computer interaction with artificial intelligence to solve the problems of low physical fitness among university students (T. Wang & Park, 2021). Moreover, the study S11 focused on a recommendation system that can automatically generate personalized sports training programs to effectively improve the strength level of experimenters while maintaining their body shape (Chen & Tian, 2024). In Study S12, they propose a personalized marathon training plan monitored in real time with various physiological data, including heart rate variability and lactate threshold. Finally, it performs a training evaluation analysis to issue intelligent recommendations (GUO, 2025). In agreement, the study S14 creates a sports plan recommendation model. It analyzes the characteristics and challenges of user data in online sports software. These data include personal information, exercise habits, exercise goals, health data, etc. These data are characterized by their diversity, comprehensiveness, and dynamism to improve users' quality of life (Deepak & Anguraj, 2023). Studies S16 and S18 implement a model based on the training history of users with similar conditions to recommend the next activity for each individual. As more data is collected, the method is strengthened with more patterns learned through machine learning based on computational cognitive architecture, thereby improving the probability of the individual successfully completing the planned activity (Mahyari & Pirolli, 2021; Mahyari et al., 2022). Finally, S20 focused on an intelligent recommendation system using comparative analysis based on thermal radiation image processing technology to effectively identify athletes' fatigue status, muscle activity, and heat distribution, in order to provide them with more accurate exercise guidance. In team sports, coaches can understand each athlete's physical consumption and fatigue level in real time through thermal radiation images, allowing for more rational staff rotation and tactical adjustments. In individual sports, athletes can adjust their training intensity and rest time based on information from thermal radiation images to achieve optimal results (Lijuan et al., 2025), and S21 presented a recommendation system to improve physical exercise and solve obesity and poor health problems faced by people with fast-paced and stressful lifestyles (Liu & Cao, 2024), together with S22 proposes a case-based reasoning system to generate personalized training recommendations for marathon runners based on their training history and that of similar runners with similar running goals.

G1 - Subgroup 2 - Nutrition. - S1 focused on dietary recommendations suggesting Indian nutritional foods to improve endurance performance and recovery in female athletes (Ramaraj et al., 2024). Study S2 focused on a group-based food recommendation system to reduce disease (Rostami et al., 2024). Study S8 focused on a post-workout meal recommendation system to improve athletic performance and promote muscle recovery. Food product recommendations based on previous activities can be an effective tool for personalizing recommendations for each user (Kumar & Sahani, 2023). Study S15 focused on a personalized recommendation model for carbohydrate and protein supplement intake to improve performance based on endurance exercise (X. Wang et al., 2023).

G1 - Subgroup 3 - Motivation. - S17 focuses on a motivational recommendation system for athletes to improve daily physical activities (Deepak & Anguraj, 2023). Study S19 proposes an effective motivational recommendation system to provide a motivational boost when they reach a level of mental saturation after an athletic routine (Deepak et al., 2022).

G1 - Subgroup 4 - Injuries. Study S4 created an intelligent system to assess, predict, and detect sports injuries to improve athletes' performance, as they may be off the field for an extended period (Meng & Qiao, 2023).

G1 - Subgroup 5 - Education. Finally, S13 proposes a course recommendation system for personalized technical and vocational training programs (Hassan et al., 2024).

The second group, G2, refers to the metrics used in SR. This group includes studies that used methods and techniques to evaluate machine learning models to improve SR in amateur cyclists.

Study S1, S3, S5, S8, S13, S14, S16, S18, S21, S22 applied metrics accuracy, F1 Score, precision, recall, Receiver Operating Characteristic Area Under the Curve (ROC AUC), and Matthews Correlation Coefficient (MCC) to measure the performance of the presented RS model. In study S4, they used accuracy, sensitivity, and specificity to measure the performance of the created model. Study S8 used precision, recall, and mean average precision (MAP) to evaluate their RS model. Study S2 focused on precision, recall, normalized discounted cumulative gain (NDCG), and health to measure and evaluate the performance of the RS model. Study S17 used the Wilcoxon statistical measure metrics to measure the proposed RS model. Study S6 focused on hit rate (HR@5) and mean precision (MP) metrics to evaluate the performance of the proposed model. Study S15 focused on Mean Absolute Error (MAE) and Gradient Boosted Regression Trees (GBRT). Finally, studies S7, S9, S10, S11, S12, S19, and S20 did not report which metrics they used in their work proposal text.

Regarding Research Question 2 (RQ2), we sought to investigate the benefits of using RS for amateur cyclists. Our second research question was RQ2: What are the benefits of using the Recommender System for amateur cyclists? This question focuses on whether the analyzed studies provide evidence of the benefits generated by using RS for amateur cyclists. To analyze this question, we considered the benefits reported by the authors for using RS for amateur cyclists. The benefits were categorized as athletic performance, health, and muscle recovery. Figure 9 shows the classification of the studies' benefits according to these criteria in percentage terms.

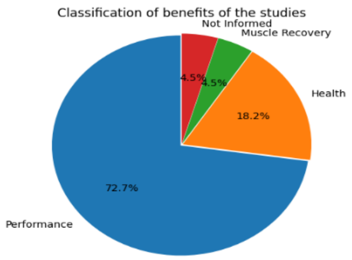

Fig. 9. Classification of benefits of the studies

Figure 9 shows that 16 studies (72.7%) (S1, S3, S4, S5, S6, S7, S9, S10, S11, S12, S14, S15, S17, S18, S19, S22) presented SRs with benefits for improving athletic performance. Approximately four studies (18.2%) (S2, S16, S20, S21) focused on SRs of foods and sports plans to improve physical exercise, solve obesity and poor health problems faced by people with a fast-paced and stressful lifestyle. Only one study (4.5%) (S8) presented a post-workout meal recommendation system to improve athletic performance and promote muscle recovery. Finally, one study (4.5%) (S13) did not report any benefits in the text of its work proposal.

Regarding Research Question 3 (RQ3): After investigating the goals and benefits of RS use, we also sought to investigate the difficulties encountered in its use. Therefore, our third definitive research question was: RQ3. What are the difficulties encountered in using RS among amateur cyclists? This question focuses on whether the analyzed studies provide evidence of the difficulties generated by the use of RS among amateur cyclists. To analyze this question, we posed the difficulties identified by the authors regarding the use of RS among amateur cyclists. Difficulties were classified as positive (some difficulty in use) or negative (no difficulty in use), and with EE (with empirical evaluation if a research method indicated in item 5 of the extraction form was used) or without EE (without empirical evaluation if no research method was used and it represents only the author's opinion). If no evidence of difficulties was indicated in the study, it was classified as "No evidence." Figure 10 shows the classification of study difficulties according to these criteria in percentage terms. It can be observed that 13 studies (59.1%) presented positive evidence of difficulties, 6 of which were positive with EE (27.3%) (S10, S16, S18, S19, S20, S22) and seven were positive without EE (31.8%) (S1, S2, S5, S7, S8, S9, S13). Less than half of the studies (40.9%) (S3, S4, S6, S11, S12, S14, S15, S17, S21) did not indicate evidence of difficulties. No article presented negative evidence of difficulty.

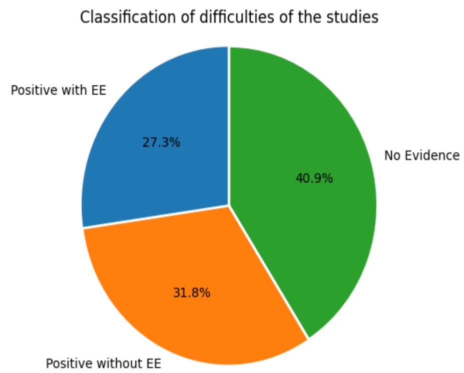

Fig. 10. Classification of difficulties of the studies.

TABLE VI		
DIFFICULTIES RESULTS		
Difficulties Evidence	Quantity	Studies
Positive with EE	6	S10, S16, S18, S19, S20, S22
Positive without EE	7	S1, S2, S5, S8, S9, S13
Negative with EE	-	-
Negative without EE	-	-
No Evidence	9	S3, S4, S6, S7, S11, S12, S14, S15, S17, S21

In order to explain the difficulties related to the EE, Table VI shows the studies by classification. In the next paragraph, the contributions are detailed.

Positive challenges with EE: Study S10 indicated that it is necessary to continuously optimize the RS in terms of function, performance, operability, compatibility, and security; as the system accumulates data, it is necessary to consider the use of big data, data mining, and other advanced technologies to establish a data warehouse for statistical analysis of various data from the sports training of college students, thus providing data support for school sports management decision-making (T. Wang & Park, 2021). Study S16 indicated that the limitations of small users or when systems are implemented without prior user history may fail in model performance (Mahyari & Pirolli, 2021). Study S18 specifies that the main challenge in developing recommender systems is personalizing them, especially for new users, when the training data set is unavailable. It employs an expert to condition the recommendation system (Mahyari et al., 2022). Study S19 indicated that the developed model does not work for wearable devices (Deepak et al., 2022). Study S20 highlighted the limitations in infrastructure and equipment for university sports activities to generate a sports community (Lijuan et al., 2025). Study S22 indicated the importance of measuring the average or maximum heart rate during training to accurately assess training results (Feely et al., 2023).

The positive challenges without EE: Study S1 confirmed that the system will be able to recommend to athletes the recommended food intake over a specific period and measure their physical results in real time (Ramaraj et al., 2024). Study S2 indicated that it uses a single user's rating for group prediction, which could affect the results (Rostami et al., 2024). Study S5 indicated that the model's weakness is its use of protocol-dependent load attributes, and therefore suggested addressing this deficiency (Chikov et al., 2022). Study S8 clarified that by including more data sources, such as user reviews, professional comments, and social media trends, the system could be better able to track changing dietary preferences and meet specific dietary demands (Kumar & Sahani, 2023). Study S9 indicated issues with the feasibility of cyclists' suggestions, primarily related to activity frequency and feasibility (Demosthenous et al., 2022). Finally, Study S13 confirmed that incorporating additional factors, such as cultural implications, social norms, values, educational systems, access to technology, economic conditions, and institutional infrastructure in different countries, could further improve the effectiveness of the presented model.

Thus, in an attempt to answer our third research question, we can observe that 59.1% of the studies presented positive evidence of difficulties, such as: The incorporation of big data, data mining; previous user history for model performance especially for new users when the training dataset does not exist; importance of taking the average or maximum heart rate for the sports performance model; use of load-dependent attributes; more data sources, such as user reviews, professional comments and social media trends.

4.3. Discussion

The studies (S3, S5, S6, S12, S14, S22) propose improving athletes' physical condition and health in athletics using RSs for sports plans. They agree that demotivation in endurance exercise is multifaceted, encompassing physiological fatigue, biochemical changes, psychological factors, and external barriers. Addressing these problems through pleasurable exercise modalities, structured support systems, and an understanding of the psychobiological aspects of perceived exertion can help mitigate demotivation and improve adherence to endurance exercise routines, as presented in studies (S17, S19). The RSs motivate athletes in their training routines, a topic that has been little explored in amateur cycling, a sport with similar characteristics to athletics, in this sense, the studies (S1, S2, S8, S15) propose that the Fitness Center RS approach to nutrition plays a crucial role in improving fitness and athletic performance, making it a key focus for fitness centers (Schroeder, 2022). In short, gyms should prioritize evidence-based nutritional guidance, ongoing education, and personalized strategies to support the health and performance of their users. Adequate hydration, balanced macronutrient intake, and adequate micronutrient consumption are key components of an effective nutrition plan for athletes of any discipline. The studies (S10, S11, S16, S18, S20, S21) propose that the Fitness Center RS approach to training plans improves cardiovascular fitness, muscular strength, endurance, flexibility, and balance, and cross-training routines to optimize amateur cycling (Ben Waer et al., 2024).

The previous cited author present that each discipline have as individual computational challenge, for instance: i) Cycling is focus on "Accurate prediction of Metabolic Cost (Gross Efficiency) and Acute Nutritional Failure (GI distress risk) from environmental sensor data" (Gough & Sparks, 2024), and, running due to the weight-bearing nature of the sport is focus on "Prediction of Tissue-Specific Injury (e.g., stress fractures) based on the integration of training volume with Gait Dynamics (Ground Reaction Forces, GCT)", however the open challenge for different sport that will faced is the Quantifying and predicting the Interaction of Perception of Effort and Potential Motivation (PBM), therefore the RSs are presented as alternative for resolve it, hence the RSs allows to personalized athletic performance has become a major thrust, to dynamically adjust training load, optimize recovery, and fine-tune nutritional intake to maximize physiological adaptation and prevent training-related pathologies such as overtraining or injury.

5. CONCLUSION

In this article, we conducted a systematic review (SLR) to investigate and understand how the concept of SR has been used among amateur cyclists, providing an overview of machine learning models and metrics for using SR with amateur cyclists. Our objective was to answer three research questions, which addressed the main objectives of using SR with amateur cyclists, the benefits of using SR, and the challenges encountered throughout the process. In this sense, this article presents the full development of the systematic review protocol (planning stage) as well as its execution process. As a result, we analyze the state-of-the-art of SR use among amateur cyclists. Throughout this SLR, 181 studies were obtained, of which only 22 were selected for analysis due to the defined selection criteria.

After analyzing these studies, we conclude that the number of studies shows a slight upward trend over the years. Most of the studies come from journals (68.2%) and conferences (31.8%). The authors' main countries of origin are China, India, and the USA. The primary context of application of the studies is academia, accounting for approximately 50% of the studies. Most studies present formal research methods (68.1%), but among those that do, the main research method used is the case study (54.5%), and a minimal number of studies use controlled experiments (13.6%). The predominant educational level is higher education (university), accounting for 89.5% of the studies.

Regarding the research questions, we conclude that there is a wide variety of uses of RS in amateur cyclists. However, we found that the main objectives of its use are studies related to sports performance (80%) and studies aimed at improving health (10%). Most studies present metrics such as accuracy, F1 Score, precision, and recall (45%) to evaluate the performance of the submitted works.

Furthermore, approximately 59% of the studies presented positive evidence of benefits, such as: Positive with EE (27.3%): The incorporation of big data and data mining; the user's prior history for model performance, especially for new users when a training dataset is unavailable; the importance of taking the average or maximum heart rate for the sports performance model. However, approximately 31.8% of the studies presented positive evidence of pitfalls, such as: Positive without EE (31.8%): The use of load-dependent attributes; more data sources, such as user reviews, professional comments, and social media trends, are limitations that could be improved in RS.

The results obtained in this systematic literature review are very important for the computer science and sports communities, as they present an overview of the main published studies on how RS is used with amateur cyclists.

As future work, we intend developing adaptive RSs architectures capable of dynamically adjusting their modeling complexity, such us, setting to be able to transition seamlessly from the high-dimensional, power-centric models necessary during a cycling training block to the injury-risk, biomechanically-focused models required during a running segment, while maintaining a consistent and reliable psychobiological state model that underpins all performance recommendations. The RSs architecture will tackle the cold-start problems in RS for new cyclists, based on the individual profile, as well as approaches of collaborative filtering with similar profiles, the integration is planned use a mobile APP for the capture of data with wearable devices.

Author Contributions: Conceptualizing: J.P.V., J.M, R.S.L; methodology, J.P.V, N.S.R, J.M. R.S.L; formal analysis, J.P.V, N.S.R, J.M. R.S.L; investigation, J.P.V, N.S.R, J.M, R.S.L, N.L.R; data curation, J.P.V, N.S.R, J.M. R.S.L, N.L.R; writing—original draft preparation, J.P.V; writing—review and editing, J.P.V, N.S.R, J.M. R.S.L,

N.L.R; visualization, J.P.V, J.M, R.S.L, N.S.R, N.L.R; supervision, J.M, R.S.L; project administration, J.P.V. All authors have read and agreed to the published version of the manuscript

REFERENCES

Andric, M., Ivanova, I., & Ricci, F. (2021). Climbing Route Difficulty Grade Prediction and Explanation. ACM Int. Conf. Proc. Ser., 285-292. Scopus. https://doi.org/10.1145/3486622.3493932

Breyer, B., Cummings, C., Hoang, K., & Ng, V. (2025). Greenspace moderates heat avoidance in physical activity during extreme heat: Evidence from Strava data in Houston, Texas. Journal of Transport & Health, 43, 102058. https://doi.org/10.1016/j.jth.2025.102058

Chen, Y., & Tian, X. (2024). Generation Algorithm of Personalized Sports Training Program Based on Sensor Data. IEEE Int. Conf. Electr., Autom. Comput. Eng., ICEACE, 1291-1296. Scopus. https://doi.org/10.1109/ICEACE63551.2024.10898477

Chikov, A. E., Pavlov, E. A., Egorov, N. A., Medvedev, D. S., Chikova, S. N., & Drobintsev, P. D. (2022). ARTIFICIAL INTELLIGENCE MODELING OF PHYSIOLOGICAL PARAMETERS AT ANAEROBIC THRESHOLD. Human Sport Medicine, 22, 46-53. Scopus. https://doi.org/10.14529/hsm22s206

Deepak, V., & Anguraj, D. K. (2023). Provisioning an Efficient Recommender System to Measure the Players Activities Using Machine Learning Approaches. International Journal of Reliability, Quality and Safety Engineering, 30(4). Scopus. https://doi.org/10.1142/S021853932350016X

Deepak, V., Anguraj, D. K., & Mantha, S. S. (2022). Recurrent neural network based recommendation system for marathoner's motivation. International Journal of System Assurance Engineering and Management. Scopus. https://doi.org/10.1007/s13198-022-01700-7

Deepak, V., Anguraj, D. K., & Mantha, S. S. (2023). An efficient recommendation system for athletic performance optimization by enriched grey wolf optimization. Personal and Ubiquitous Computing, 27(3), 1015-1026. Scopus. https://doi.org/10.1007/s00779-022-01680-2

Demosthenous, G., Kyriakou, M., & Vassiliades, V. (2022). Deep reinforcement learning for improving competitive cycling performance. Expert Systems with Applications, 203. Scopus. https://doi.org/10.1016/j.eswa.2022.117311

Feely, C., Caulfield, B., Lawlor, A., Smyth, B., & ACM. (2023). Modelling the Training Practices of Recreational Marathon Runners to Make Personalised Training Recommendations (WOS:001051715400019). 183-193. https://doi.org/10.1145/3565472.3592952

García-Peñalvo, F. J. (2022). Desarrollo de estados de la cuestión robustos: Revisiones Sistemáticas de Literatura. Education in the Knowledge Society (EKS), 23, e28600. https://doi.org/10.14201/eks.28600

GUO, L. (2025). INTELLIGENT OPTIMIZATION AND RECOMMENDATION SYSTEM DESIGN FOR PERSONALIZED TRAINING PROGRAMS FOR MARATHON ATHLETES BASED ON MACHINE LEARNING. Scalable Computing, 26(2), 864-870. Scopus. https://doi.org/10.12694/scpe.v26i2.4047

Hassan, R. H., Hassan, M. T., Sameem, M. S. I., & Rafique, M. A. (2024). Personality-Aware Course Recommender System Using Deep Learning for Technical and Vocational Education and Training. Information, 15(12), 803. https://doi.org/10.3390/info15120803

Kumar, S., & Sahani, S. K. (2023). Comparative study on post-workout meal recommender systems using machine learning algorithms. En Sharma V., Sinha J., Chandraprabha M., Kumar S., & Rana A.K. (Eds.), Proc. - IEEE Int. Conf. Adv. Comput., Commun. Control Netw., ICAC3N (pp. 378-384). Institute of Electrical and Electronics Engineers Inc.; Scopus. https://doi.org/10.1109/ICAC3N60023.2023.10541768

Lee, M. C. M., Scheepers, H., Lui, A. K. H., & Ngai, E. W. T. (2023). The implementation of artificial intelligence in organizations: A systematic literature review. Information & Management, 60(5), 103816. https://doi.org/10.1016/j.im.2023.103816

Lijuan, S., Jun, Z., & Qiao, L. (2025). Research on fine guidance of sports activities based on wearable IoT devices and thermal radiation image processing. Thermal Science and Engineering Progress, 59. Scopus. https://doi.org/10.1016/j.tsep.2025.103352

Liu, Y., & Cao, S. (2024). The analysis of aerobics intelligent fitness system for neurorobotics based on big data and machine learning. Heliyon, 10(12). Scopus. https://doi.org/10.1016/j.heliyon.2024.e33191

Mahyari, A., & Pirolli, P. (2021). Physical Exercise Recommendation and Success Prediction Using Interconnected Recurrent Neural Networks. En Ahamed S.I., Atukorala N., Chang C.K., Damiani E., De Pietro G., Liu L., Wang Z., Zhang J., & Zulkernine F. (Eds.), Proc. - IEEE Int. Conf. Digit. Health, ICDH (pp. 148-153). Institute of Electrical and Electronics Engineers Inc.; Scopus.

https://doi.org/10.1109/ICDH52753.2021.00027

Mahyari, A., Pirolli, P., & Leblanc, J. A. (2022). Real-Time Learning from an Expert in Deep Recommendation Systems with Application to mHealth for Physical Exercises. IEEE Journal of Biomedical and Health Informatics, 26(8), 4281-4290. Scopus. https://doi.org/10.1109/JBHI.2022.3167314

Mendez, D., Avgeriou, P., Kalinowski, M., & Ali, N. B. (Eds.). (2024). Handbook on Teaching Empirical Software Engineering. Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-71769-7

Meng, L., & Qiao, E. (2023). Analysis and design of dual-feature fusion neural network for sports injury estimation model. Neural Computing and Applications, 35(20), 14627-14639. https://doi.org/10.1007/s00521-021-06151-y

Ramaraj, K., Narayan, V., Dhivyaprabha, T. T., & Subashini, P. (2024). A healthy nutrition suggestion model for indian women sports players & active youth using long short-term memory. Internet Technology Letters, 7(5). Scopus. https://doi.org/10.1002/itl2.452

Rostami, M., Berahmand, K., Forouzandeh, S., Ahmadian, S., Farrahi, V., & Oussalah, M. (2024). A novel healthy food recommendation to user groups based on a deep social community detection approach. Neurocomputing, 576, 127326. https://doi.org/10.1016/j.neucom.2024.127326

Shull, F., Singer, J., & Sjøberg, D. I. K. (Eds.). (2008). Guide to Advanced Empirical Software Engineering. Springer London. https://doi.org/10.1007/978-1-84800-044-5

Thyago, T., Seiji, I., Bittencourt, I. I., & Lu, Y. (2021). Collective Intelligence in Online Educational Technologies: A Systematic Literature Review. IEEE Access. https://doi.org/10.1109/TLT.2021.3073559

Wang, T., & Park, J. (2021). Design and Implementation of Intelligent Sports Training System for College Students' Mental Health Education. Frontiers in Psychology, 12. Scopus. https://doi.org/10.3389/fpsyg.2021.634978

Wang, X., Li, Z., & Wu, H. (2023). Personalized Recommendation Method of «Carbohydrate-Protein» Supplement Based on Machine Learning and Enumeration Method. IEEE Access, 11, 100573-100586. Scopus. https://doi.org/10.1109/ACCESS.2023.3314699

Wen, Y., Liu, B., Li, Y., Chen, X., & Xu, Y. (2025). The spatial heterogeneity effects of street environmental factors on the preference for sports and leisure cycling paths across different street types. BMC Public Health, 25(1), 621. https://doi.org/10.1186/s12889-025-21717-4

Yang, Y., & Zhao, Y. (2025). Personalized Sports Health Recommendation System Assisted by Q-Learning Algorithm. International Journal of Human–Computer Interaction, 41(4), 1889-1901. https://doi.org/10.1080/10447318.2023.2295693

Zhang, Y. (2025). Athlete Physical Fitness Evaluation and Training Recommendation System Based on Machine Learning. IEEE Int. Conf. Electron., Energy Syst. Power Eng., EESPE, 1725-1731. Scopus. https://doi.org/10.1109/EESPE63401.2025.10986889

Zhao, X. (2024). Personalized Health Plan Recommendation Algorithm Based on User Data of Network Sports Software. Proc. - Int. Conf. Data Anal., Comput. Artif. Intell., ICDACAI, 43-47. Scopus. https://doi.org/10.1109/ICDACAI65086.2024.00015

Zientarska-Kayko, A., E. Brown, S., S. Ireland, T., & Agnes Zientarska-Kayko, Sarah E. Brown & Timothy S. Ireland. (2024). Power of Zotero: Case Studies from the University of Waterloo. Journal of Electronic Resources Librarianship. https://doi.org/10.1080/1941126X.2024.2417125