INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
Online ISSN 1841-9844, ISSN-L 1841-9836, Volume: 19, Issue: 4, Month: August, Year: 2024
Article Number: 6498, https://doi.org/10.15837/ijccc.2024.4.6498

CCC Publications

A Data-Driven Assessment Model for Metaverse Maturity

Mincong Tang, Jie Cao, Zixiang Fan, Dalin Zhang, Ionut Pandelica

Mincong Tang
1.Xuzhou University of Technology, China
2.Industrial University of Ho Chi Minh City, Vietnam
*Corresponding author: mincong@bjtu.edu.cn

Jie Cao
Xuzhou University of Technology, China
cj@nuist.edu.cn

Zixiang Fan
Beijing Union University, China
*Corresponding author: sftzixiang@buu.edu.cn

Dalin Zhang
Beijing Jiaotong University, China
dalin@bjtu.edu.cn

Ionut Pandelica
Bucharest University of Economic Studies
Faculty of International Economic Relations
Romania
ionut.pandelica@rei.ase.ro

Abstract

The rapid development of the metaverse has sparked extensive discussion on how to estimate its development maturity using quantifiable indicators, which can offer an assessment framework for governing the metaverse. Currently, the measurable methods for assessing the maturity of the metaverse are still in the early stages. Data-driven approaches, which depend on the collection, analysis, and interpretation of large volumes of data to guide decisions and actions, are becoming more important. This paper proposes a data-driven approach to assess the maturity of the metaverse based on K-means-AdaBoost. This method automatically updates the indicator weights based on the knowledge acquired from the model, thereby significantly enhancing the accuracy of model predictions. Our approach assesses the maturity of metaverse systems through a thorough analysis of metaverse data and provides strategic guidance for their development.

Keywords: Metaverse, Data Driven, Maturity Assessment, K-Means.

1 Introduction

Historically, the term 'metaverse' was first coined by Neal Stephenson in his 1992 science fiction novel "Snow Crash," where it was defined as a vast virtual environment parallel to the physical world, and users interacted through digital avatars. From a technical point of view, it represents the convergence of Augmented Reality (AR), Virtual Reality (VR), 3D holography, and social media platforms, providing users with an immersive experience that blends digital and physical realities [1].

The metaverse was depicted as a social digital space emerging after the completion of digital transformation, which represents a development stage in the digital economy [2]. From a cultural and sociological perspective, the 2022 Metaverse Development Research Report from Tsinghua University considered the metaverse as an embodiment of digitization and virtualization trends at the sociocultural level, reflecting people's social interactions and cultural expressions in virtual space[1]. Wang et al. defined it as a comprehensive, immersive, transcendent, and self-sustaining virtual shared space that integrates the physical, human, and digital realms, symbolizing the evolution of the next generation of the internet[3]. The metaverse is also described by Lee et al. as a multi-user interactive environment, emphasizing the significance of users when experiencing various activities and social interactions through digital avatars[4]. The metaverse is seen as an emerging market space involving transactions of virtual goods and services in business, showing the influential impact of new economic and business models[5].The definition of the metaverse not only shows a technological innovation but also reflects the profound changes in social, cultural, and economic domains.

Up to now, metaverse technology has made a significant progress. However, the methods for evaluating the maturity of the metaverse are still quite scarce. A promising development of the metaverse needs a global, systematic, scientific development framework and assessment system, which could promote the growth and finally foster its integration and interaction with the real world[5]. Previous research has been proposed to assess the maturity of metaverse systems with a maturity model, which is constructed by multiple factors. The maturity model was segmented into different levels, each with a corresponding set of evaluation indicators. By weighting and averaging the indicators at each level, an overall score is calculated, representing the overall maturity rating of metaverse technology in the transportation area. While the method has proven to be effective, but its scope is overly narrow, and is always constrained by specific assumptions and methodological limitations. And the weights of the indicators in the previous research were determined only by experts, which causes blindness and limitations for prediction. In order to solve the above problems, this paper proposes a data-driven approach for evaluating metaverse maturity, aiming to fill the gap in current assessment methods. And in the data-driven approach, we propose a method called K-means-AdaBoost, which automatically updates the indicator weights according to the knowledge learnt from the model, and greatly improves the accuracy of the model prediction.

This paper begins with an introduction about the research background in Chapter 1. Chapter 2 illustrates on the related work concerning the maturity assessment methods for the metaverse. Chapter 3 thoroughly constructs a metaverse evaluation indicator system and proposes a data-driven method for evaluating the maturity of metaverse systems. In Chapter 4, through case studies and empirical research, the specific application of the data-driven method in metaverse evaluation is demonstrated. This includes the selection of cases, data processing, model validation, and the discussion of results. Finally, Chapter 5 concludes the main findings, limitations and future works of this paper.

2 Literature review

2.1 The metaverse evaluation methods

The metaverse can be described as a digital ecosystem which is based on augmented reality, virtual reality, and mixed reality technologies[6]. It constructs a three-dimensional space as a digital extension that blends and interacts with it. And the blend and interactivity are core characteristics of the metaverse[5]. Researchers generally agree that the essence of metaverse technology lies in digital twinning technology[5].This means the metaverse is not just a virtual digital space but also includes mechanisms and technologies for interaction with the real world[7]. Users can be as digital identities in the metaverse and interact with the virtual environment, other users, or the physical world. A robust metaverse system encompasses various cutting-edge technologies for each playing a crucial role. For instance, Augmented Reality (AR) technology can provide richer virtual information and navigation, helping users better integrate in the intersection of the virtual and real worlds. Virtual Reality (VR) technology offers highly immersive experiences for the metaverse. The integrated application of these technologies creates a diversified and highly interactive metaverse environment, where users can experience, create, and share content.

The metaverse is a digitalized, enhanced, multidimensional space interacting with the real world[3]. In order to more accurately master the developing state of the metaverse with applied a variety of complex technologies, many researchers have tended to evaluate the maturity of the metaverse with their approaches. The development of the metaverse demands the continuous innovation and improvement, and thus it also requires regular assessments to guide its developmental direction[7]. The diversity of assessment methods shows the depth and breadth of the metaverse field. Zhao Xing et al. proposed the use of the metaverse index, duration, and population as quantitative metrics[2]. The Metaverse Report (2021 2022) produced by Fudan University proposed several core dimensions for the metaverse, including computing power, responsiveness, realism, immersion, interactivity, user autonomy, digital property protection, and digital currency payments. And Zainab et al. proposed that there are many advanced technologies such as virtual reality, augmented reality, artificial intelligence, blockchain, etc applied in the metaverse[8]. Although there are various strengths in different dimensions and different components, but they all contribute to the metaverse concept, leading to the notion of a 'metaverse rate'.

Weinberger et al. assessed the maturity of the metaverse with a metaverse maturity model[9], specifically including 8 core attributes and 5 maturity levels. Additionally, Deveci et al. proposed a model-based assessment method[10] to evaluate the application of metaverse technology in the transportation area. The research by Lee LH[4] categorized the development of the metaverse into three stages: digital twins, digital natives, and the coexistence of physical and virtual realities, revealing the gradual maturation of metaverse technology.

The above methods demonstrate the diversity and complexity of metaverse maturity assessments, and covered not only the technical aspects but also social, economic, and cultural influences, which provide a comprehensive perspective about the development of the metaverse. Although each assessment method has its strength, all of them are also constrained by their specific assumptions, scope, and methodological limitations. In contrast, a data-driven approach is a decision-making and operational method based on the collection, analysis, and interpretation of vast amounts of data to guide decisions and actions. This approach considers data as the most critical basis for decision-making, and through in-depth analysis, it can reveal hidden patterns and insights to derive valuable knowledge. In a data-driven environment, organizations and individuals can optimize business processes, improve products and services, and achieve higher business goals with the data.

The data-driven methods spans across various industries for applications, providing more accurate and effective support for decision-making. For instance, Jin Yao [11] studied how data analysis could enhance the user experience in mobile service halls in the telecommunications sector. For the military and aerospace sectors, Song Yining[12] explored data-based methods for space situational awareness to monitor and predict changes in space environments more accurately. Li Yanping's research[13] focused on improving the quality and efficiency of library services by using data-driven method for intelligent library management. Furthermore, Wu Hao[14] studied traffic signal control systems with data-driven method, which can be used to optimize traffic flow through real-time data analysis, and enhance the use of road efficiency. These application examples highlight the versatility and practicality of data-driven methods, reflecting their widespread use and profound impact in various fields.

Considering the effectiveness of the data-driven method, this paper proposes a metaverse maturity assessment method based on data-driven. This method uses data as the primary decision-making basis, and through in-depth analysis of metaverse data, the maturity level of metaverse systems was finally evaluated , which could provide guidance for the development of metaverse.

2.2 Metaverse Evaluation Indicators

The metaverse evaluation indicator system plays a crucial role in the development of the metaverse. Since the development of the metaverse requires comprehensive consideration of technology, society, economy, and other aspects[5], it's essential to establish a comprehensive evaluation indicator system for the metaverse. Through applying the indicator system, we can assess the metaverse's development in terms of technological level, social impact, economic benefits, etc., and get a better support for guiding the development of the metaverse.

Zhao Xing et al. proposed three basic evaluation indicators for the metaverse in their paper: the metaverse index, metaverse duration, and metaverse population[2]. These three indicators represent the space, time, and development status of the metaverse, as detailed in Table 1:

Table 1: The quantifiable metrics for measuring metaverse development

Indicator Name	Measurement Method	Measurement Meaning	Historical Milestone
Metaverse Index	The ratio of consumer-end users owning (or frequently using) metaverse devices such as VR/AR.	This is a metaverse hardware penetration rate indicator.	The metaverse becomes a significant part of daily life.
Metaverse Duration	The proportion of time spent in metaverse applications within 24 hours	This is an indicator for metaverse applications.	Human reliance on the metaverse surpasses that of the physical universe.
Metaverse Population	The ratio between the number of active virtual populations (digital beings, avatars) to the biological population	This is an indicator of the metaverse's socialization process.	Silicon-based life becomes the mainstream life form.

2.3 Metaverse Evaluation Methods

Currently, some researchers have proposed a number of metaverse assessment methods. Pamucar et al. proposed a novel Rough Aczel–Alsa (RAA) function and the Ordinal Priority Approach (OPA) method to assessment the metaverses in the public transportation field[10]. Roverto et al. evaluate metaverse ecosystems by using a multilayer network, which explores the transition from physical reality[15] to metaverse by analysing the connections and interdependencies between the physical world, the network and the metaverse[16]. Traditional evaluation methods encounter challenges in addressing the intricate, dynamic, and cross-platform nature of the metaverse. Machine learning provides a new approach to addressing these challenges. Kim et al. employed deep learning to identify abnormal behavior in the metaverse[17]. By integrating machine learning, big data analytics, and metaverse features, evaluation models based on machine learning can automatically identify issues, providing decision-makers with scientific and quantifiable results. These data-driven approach demonstrate robust capabilities in modeling complex systems, enabling the exploration of intricate relationships among system indicators and the development of models adaptable to diverse scenarios [18]. The data-driven approach has been demonstrated to effectively model complex systems by analyzing the intricate relationships between system metrics and constructing models that are adaptable to various situations. And the data visualization is the scientific and technical mainly about visually representing data, where this visual representation is defined as a form of abstracted information presented in a summarized manner. It encompasses various attributes and variables of the related information unit, and it is a dynamic concept with constantly expanding boundaries.

The evaluation method based on machine learning typically involves the following steps: (1) construction of an evaluation index system; (2) collection of metaverse product indicator data; (3) feature extraction; (4) model training; and (5) quantitative assessment. In this paper, we developed a model for assessing the maturity of metaverse products based on a model called K-means-AdaBoost. By learning the objective weights of indicators at each levels, it is possible to avoid the subjectivity of manual evaluations. This model can improves the accuracy and efficiency of assessing metaverse product maturity.

3 Research methodology

3.1 Metaverse Evaluation Indicator

During the exploration of the 3D virtual worlds , John et al. identified four crucial features of the metaverse: realism, ubiquity, interoperability, and scalability [6]. These four elements are deemed essential components in bringing the metaverse to fruition, serving as key criteria for evaluating its development. Currently, there is no official, credible method for assessing metaverse maturity, despite the numerous evaluation indicator systems exists in the field. This means that when evaluating the maturity of the metaverse, there is a lack of a unified and reliable standard. To effectively drive the development of the metaverse, we urgently need to establish a metaverse maturity evaluation method that comprehensively measures its maturity level in multiple aspects, including technology, society, and economy. Such an evaluation method will help to guide the industries focus on key development areas, and ensure its broader application and long-term prosperity. Drawing upon the latest research findings, we designed a multi-tiered, comprehensive evaluation indicator system for assessing the maturity of the metaverse, as presented in Table 2:

Table 2: Multilevel and Comprehensive Evaluation of Indicator System

Primary Indicator	Secondary Indicator	Tertiary Indicator
User Autonomy	Personalized Customization	Interface layout customization
		Theme and color scheme
	Freedom of Content Creation	Content publishing permissions
		Diversity of creation tools
Immersiveness	Ultimate Immersive Experience	Realism of the virtual environment
		User interaction feedback
	Virtual Reality Technology	Hardware accessibility
		Rendering technology level
	Interaction Design	Efficiency of user input
		Navigational usability
Sociability	Social Interaction Alternatives	Virtual gathering functionalities
		Sharing of social dynamics
	Social Network Integration	Third-party social media connectivity
		Data interoperability
	Community Engagement	Forum and discussion board activity
		Community-organized events
Openness	Freedom of Play	Exploration scope of the world
		Freedom in behavior choices
	Open APIs to Third Parties	Range of API functionalities
		Completeness of interface documentation
	Platform Compatibility	Operating system support
		Device adaptability
	Provision of API/SDK	Completeness of interface documentation
		Developer support services
Sustainability	Perpetual Operation	Business model innovation
		Investment return cycle
	Sustainable Development Strategy	Environmentally friendly technologies
		Implementation of social responsibilities
	Environmental Impact Consideration	Resource consumption evaluation
		Carbon footprint calculation
Rich Content Ecosystem	Vibrant Creator Economy	Creator revenue models
		Market acceptance of created content
	Diversity	Content types
		Diversity of user backgrounds
	Update Frequency	Content update cycle
		Speed of feature iteration
Comprehensive Economic System	Bridging Virtual and Real	Integration level of technology
		Consistency of user experience
	Civil Rules	User behavior guidelines
		Conflict resolution mechanisms
	Currency System Design	Currency stability
		Transaction security

The above system provides a structured approach to evaluate the maturity of the metaverse, considering a wide range of factors from user experience to economic viability, offering a robust framework for holistic assessment. In this system, user autonomy is central, and personalized customization is the primary component of this indicator, which empowers users to adjust interface layouts and select themes according to personal preferences. Such customization is crucial for long-term user engagement and satisfaction. Immersiveness is one of the key factors in the metaverse system[19]. The ultimate immersive experience encompasses various technologies and designs, including how to simulate a realistic virtual environment with high-quality visual, auditory, and even tactile feedback. The quality of this experience directly impacts user involvement in the metaverse. In terms of sociability, the ability to substitute for real-world social activities should be fully considered in the design of the metaverse., which allows users to engage in social interactions within digital spaces[20]. Openness make users can freely explore and act within the virtual world. The sustainability metric focuses on whether the metaverse platform can operate long-term and maintain competitiveness. The richness of the content ecosystem is manifested through a vibrant creator economy, where creator revenue models and the market acceptance of created content are key metrics. This relates not only to whether creators can receive fair compensation for their work but also affects the maintenance of content diversity on the metaverse platform. A comprehensive economic system contains about how to establish a stable, secure monetary system, including designing mechanisms that support both virtual and real economic activities and establishing civil rules to maintain order and fairness.

Overall, the above comprehensive indicators cover not only technological aspects but also include multiple dimensions such as social, economic, and cultural, providing a thorough evaluation framework for the development of the metaverse[21]. Through such an evaluation system, a better understanding of the current state of metaverse platforms can be achieved, enabling the formulation of strategies to promote their development, while also providing clear reference standards for users and investors to make more informed decisions.

3.2 Data-driven metaverse evaluation method

Data-driven refers to a method of extracting information and knowledge from data and using it for decision-making and optimization through techniques such as data analysis and machine learning. Data-driven methods tend to be more closely aligned with reality and better reflect the actual laws governing objective phenomena. Most machine learning and deep learning data-driven models differ significantly from statistical models, as they no longer require a focus on assumptions such as data distribution or probability. Instead, the emphasis is on the model's fit to the training data and the prediction of the test data, resulting in fewer assumptions compared to traditional mathematical and statistical learning models. Consequently, these models are easier to build. Finally, data-driven methods such as machine learning and deep learning can greatly benefit from advancements in computer technology, sensing technology, database technology, data processing technology, and cloud computing. These advancements enable the training of models based on massive amounts of data, ensuring their practicality and reasonableness. Additionally, cloud computing technology reduces the training time to meet people's expectations.

The data-driven metaverse evaluation method is an assessment approach that relies on data and indicators to evaluate the current status and future trends of metaverse development. The method objectively, accurately, and scientifically evaluates all aspects of the metaverse by collecting and analyzing relevant data in areas such as enterprise information, technology research and development data, and market data. It also involves constructing a corresponding evaluation index system.

Data-driven metaverse evaluation methods typically include the following considerations:

(1) Defining goals and metrics: Clearly defining the objectives of a data-driven metaverse is essential. This may involve providing a more realistic virtual experience, supporting more sophisticated data analysis and visualization, or enhancing user interaction. Then, identify key metrics for evaluation that are directly related to the goals and desired outcomes of the metaverse.

(2) Data quality and consistency: When evaluating a data-driven metaverse, it is important to consider the quality and consistency of the data within it. Ensure that the data in the metaverse is accurate, complete, and reliable to support a wide range of applications and analytical requirements.

(3) User experience: The success of the metaverse depends on the user experience. Evaluate user interaction, navigation, and experience in the metaverse to ensure that they can easily and efficiently utilize metaverse features.

(4) Performance and Scalability: It is important to consider the performance and scalability of the metaverse to ensure that it can effectively handle large volumes of data, users, and applications

without compromising performance.

(5) Security and privacy protection: When evaluating the data-driven metaverse, ensure that the system's security and the privacy of its users are adequately addressed. This includes secure storage and transmission of data, as well as the handling of data in compliance with privacy regulations.

(6) Evaluation of visualization and interaction tools: Evaluate the visualization and interaction tools in the metaverse to ensure they effectively convey information, facilitate analysis, and provide intuitive user interfaces.

(7) Feedback and improvement mechanisms: Implement feedback mechanisms to enable users and system administrators to report problems, suggest improvements, and ensure that these suggestions are incorporated into the system improvement plan promptly.

(8) Monitoring and analysis: Establish a monitoring system to regularly analyze the usage and performance data of the metaverse, along with user feedback, to continuously optimize and enhance the system.

The data-driven process for evaluating the metaverse is illustrated in Figure 1. In the process of assessing the maturity of the metaverse using data, we begin by acquiring data from the cloud. After completing data cleaning, preprocessing, and analysis, we extract key features.

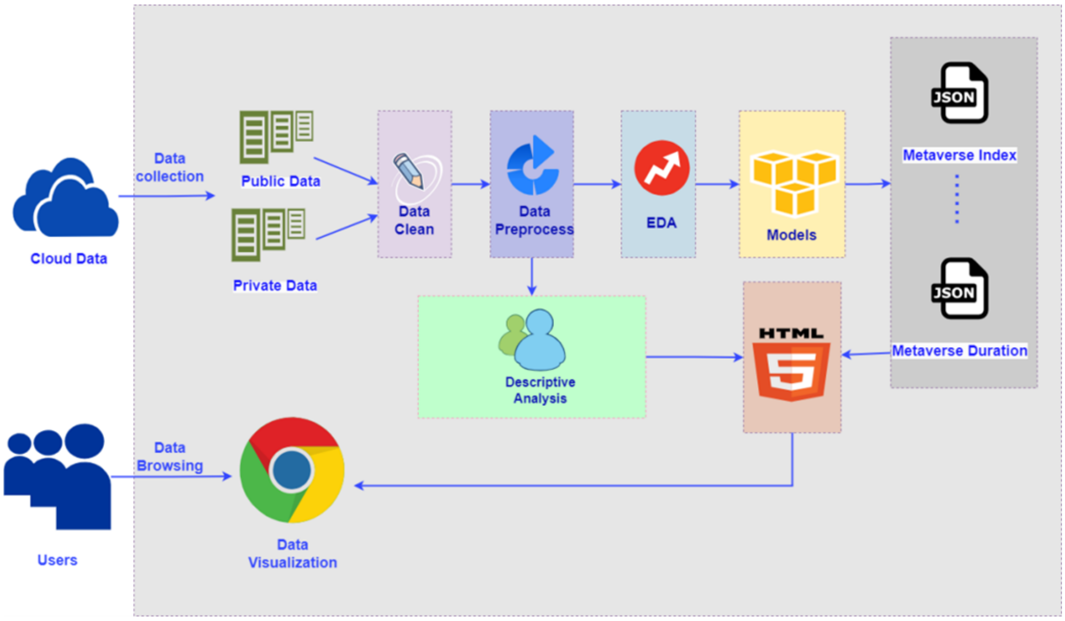

Figure 1: The data-driven evaluation process for maturity of metaverse

3.3 Metaverse Maturity Evaluation Model

3.3.1 Division of Evaluation Levels

In this study, in consideration of practical scenarios, we categorize metaverse product maturity into five levels, denoted as $R = R_1; R_2; R_3; R_4; R_5$. Here, R_1 represents the lowest level, while R_5 signifies the highest level. The division of levels is conducted using intervals within the range (0,100], and the corresponding score intervals from low to high are (0,30], (30,50], (50,70], (70,90], (90,100]. The specifics are outlined in the table below:

Table 3: Table of Corresponding Maturity Scores and Levels

Score Range	Level R	Level Description
(0,30]	Level One(R_1)	Conceptual Level
(30,50]	Level Two(R_2)	Experimental Level
(50,70]	Level Three(R_3)	Engineering Level
(70,90]	Level Four(R_4)	Market Level
(90,100]	Level Five(R_5)	Industrial Level

3.3.2 Model Overview

The maturity assessment process of the metaverse based on clustering and AdaBoost algorithms is illustrated in the figure below.

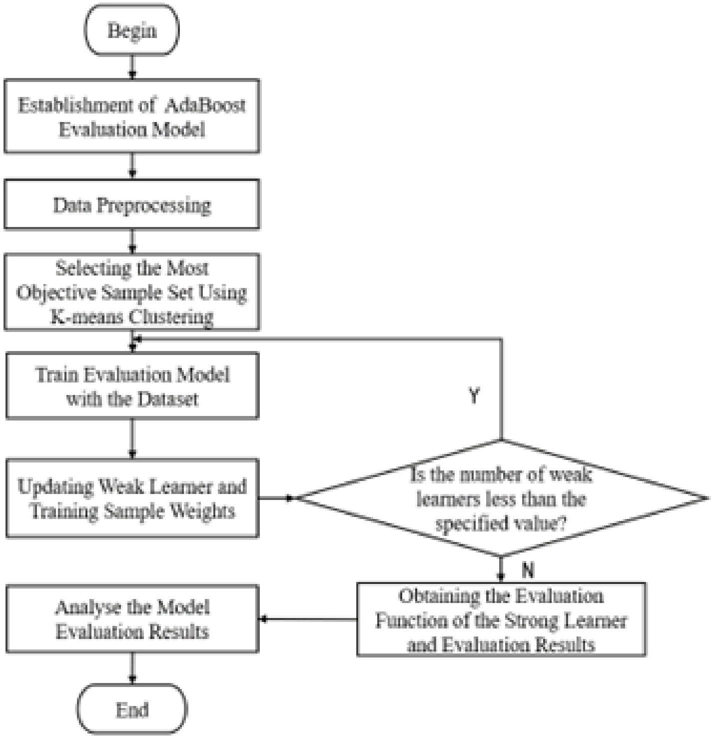

Figure 2: Process of Metaverse Maturity Assessment based on Clustering and AdaBoost Algorithm

Firstly, to reduce the impact of subjective factors, we utilize the K-means clustering method on preprocessed sample data to identify the majority of objectively assessable samples for training and testing. Subsequently, an AdaBoost model is trained using the training set samples, with a BP neural network as a weak learner. Based on user ratings for each metric and the final overall maturity score, our AdaBoost model learns and optimizes the weights for each metric. The results from the test set indicate that our model learns weights that better align with the overall maturity score labels of the test set, compared to the average weight method. Our model consists of two processes: clustering and weight learning and optimization. The following sections will provide a detailed explanation of our model.

3.3.3 K-means Clustering

Due to the subjective nature of user ratings in the evaluation system for various indicators and the final maturity assessment, we have adopted a clustering approach to identify the most objectively evaluated samples for our training and testing datasets. Clustering involves identifying the inherent properties and characteristics of unlabeled data and then grouping them into distinct subsets, with each subset forming a cluster. The K-means clustering algorithm is well-suited for large datasets compared to other clustering methods. It converges quickly and aims to partition the data into K independent clusters. The algorithm minimizes the distance between points within each cluster while maximizing the distance between clusters. The specific process is as follows:

(1) Determine the initial k cluster centers for the samples, denoted as Z_j.

(2) Calculate the Euclidean distance from each sample point x_i to the nearest cluster center Z_j. The Euclidean distance reflects the dissimilarity between categories, and samples are then reclassified based on the principle of minimum distance.

(3) Recalculate the mean for each cluster and use this mean as the new cluster center.

(4) Repeat steps 2 and 3 until each cluster center converges to a constant value.

To determine the optimal number of clusters, we use the Calinski-Harabasz (C-H) index to evaluate the clustering results. The CH index is defined as follows:

$$CH = \frac{t_B/(k-1)}{t_W/(N-k)}$$

Where $t_B = \sum_{j=1}^{k} n_j \parallel z_j - z \parallel^2$, $t_W = \sum_{j=1}^{k} \sum_{x_i \in z_j} \parallel x_i - z_j \parallel^2$, N is the total number of records, z is the mean of the entire dataset, and z_j is the mean of the j-th cluster. The CH index is positively correlated with the quality of clustering results. A higher CH value indicates tighter clusters and greater dispersion between clusters, suggesting a better clustering outcome.

3.3.4 AdaBoost Regression

AdaBoost is an ensemble learning technique that constructs a robust model by aggregating multiple weak learners. The main idea of the AdaBoost algorithm is to prioritize previously misclassified samples in each iteration. It achieves this by adjusting the weights of the samples to train new weak learners, and then combining the predictions of these weak learners through weighted summation. The AdaBoost algorithm can be outlined in three main steps: Initialization: Initialize the weight distribution D(1) for the training data, assigning equal weights to each sample. Train a weak learner using the data. Weight Adjustment: If a training sample is accurately classified by the weak learner, decrease its corresponding weight when constructing the next training set. Conversely, if a training sample is misclassified, increase its weight to give it more significance. These updated weights are then used to train the next weak learner, and the process iterates. Combining Weak Learners: After training all the weak learners, combine them to create a strong learner. The combination involves assigning higher weights (which give higher influence) to weak learners with lower classification error rates and lower weights (which give lower influence) to those with higher error rates. This ensures that weak learners with lower error rates, who are weaker, play a more significant role in the final classification function. In summary, weak learners with lower classification error rates are assigned higher weights in the final model, while those with higher error rates are assigned lower weights. This weighting strategy optimizes the overall performance of the AdaBoost model.

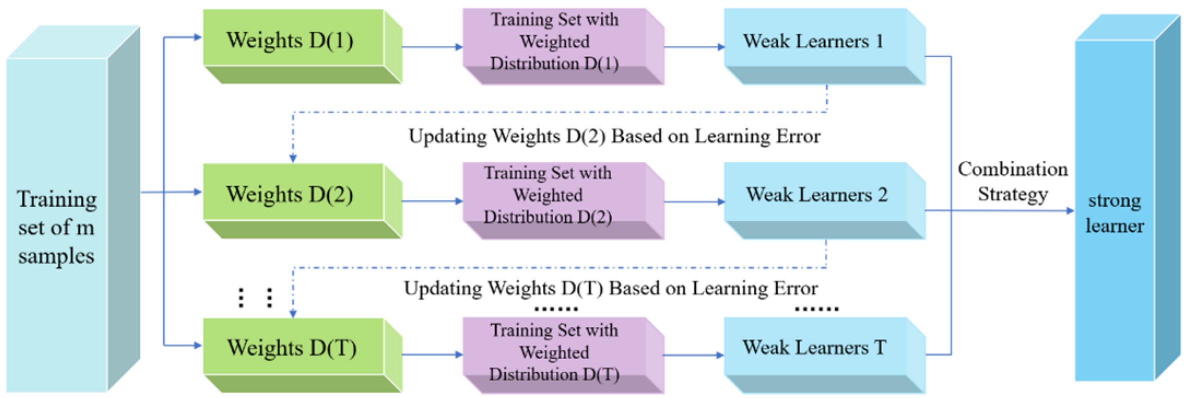

Figure 3: The Process of the AdaBoost Algorithm

We use the BP neural network as the base learner to ultimately form a strong learner. Initially, the network structure and the number of weak learners (T) are set for the BP neural network. Let's consider a dataset D composed of N samples:

$$D = \{(x_1, y_1), (x_2, y_2), ..., (x_N, y_N)\}$$

Where $x_i \in R^d$ represents the ratings of d indicators in sample i, and y_i is the overall maturity score of the sample i. The specific calculation steps are as follows: (1) Initialize the weights of training samples. The calculation process is given by:

$$D_t(i) = \frac{1}{N}$$

Where $D_t(i)$ represents the distribution of weight values for training samples, $t \in T$ denotes the $t - th$ weak learner, and $i \in N$ represents the $i - th$ training sample.

(2) Train the BP neural network using the training samples to obtain the prediction sequence $g(t)$ for the $t - th$ BP weak learner. Calculate its prediction error as:

$$e_t = \sum Dist_t(x_i)(g(t)^1 y_i)$$

Here, y represents the expected overall maturity score. (3) Calculate the weight of the $t - th$ BP weak learner based on e_t:

$$a_t = 1/2ln(\frac{1 - e_t}{e_t}), wheret = 1, 2, 3, ..., T$$

(4) Adjust the weights of training samples for the next iteration:

$$Dist_(t + 1)(x_i) = \frac{Dist_t(x_i)}{B_t} \cdot exp\left[-a_t y_i g_t(x_i)\right]$$

Here, B_t is the normalization factor, and $g_t(x_i)$ is the prediction sequence of the $t - th$ weak learner for classifying x_i. (5) Return to step 2 and iterate until T iterations are completed. (6) Output the strong learner by combining T sets of weak learning functions $f_t(x)$ according to the weight distribution to generate the strong learning function.

$$F(x) = sign\left[\sum_{t=1}^{T} a_t \cdot f_t(x)\right]$$

4 Case Analysis and Results Discussion

4.1 Data Collection and Preprocessing

In this study, data were collected through a questionnaire survey. The questionnaire was designed based on the 42 tertiary indicators, 21 secondary indicators, and 7 primary indicators outlined in Section 3.1. Each tertiary, secondary, and primary indicator corresponds to a question, and the final overall maturity score is designed as a single question. The questionnaire consists of a total of 71 questions, with each question offering six options labeled as A, B, C, D, E, and F. Each option corresponds to scores of 0, 20, 40, 60, 80, and 100, respectively. The collected sample data represents the scores for each question in the questionnaire, reflecting the evaluation indicators' scores. The study focuses on a virtual reality (VR) social platform, referred to as Platform A, as a case study. Data about this platform was gathered through an online survey. The total number of collected questionnaires was 636, with 30 invalid responses excluded.

The questionnaire data was collected and organized to obtain raw data, as shown in Table 4. Python was used to preprocess the raw data and generate input sample data, as illustrated in Table 5.

4.2 K-means Clustering Results

After preprocessing the data, K-means clustering was used to group the data. The optimal number of clusters was determined based on the Calinski-Harabasz (CH) index. The CH index values for different numbers of clusters are presented in Table 6.

According to the definition of the CH index, a higher CH index indicates better clustering effectiveness. Therefore, we selected the clustering configuration with the highest CH index value, which corresponds to 5 clusters, as the optimal number of clusters. After determining the optimal number of clusters to be 5, and taking into account the objectivity of the dataset, we choose the class with the

Table 4: Raw Data Obtained from the Questionnaire Survey

\	Primary Indicator		Secondary Indicator		Tertiary Indicator		Overall Score
Sample	**User Autonomy**	**Comprehensive Economic System**	**Personalized Customization**	**Currency System Design**	**Interface layout customization**	**Transaction security**	**Overall Score**
S1	60	80	60	80	80	80	60
S2	80	60	40	60	60	80	80
S3	60	20	40	60	60	80	100
S4	40	40	40	60	80	60	80
...	...	...	...	...	...	...	...
S603	100	100	80	40	60	60	40
S604	40	60	40	60	20	80	60
S605	40	80	40	60	20	80	60
S606	80	80	60	40	80	40	80

Table 5: Preprocessed Sample Data

\	Primary Indicator		Secondary Indicator		Tertiary Indicator		Overall Score
Sample	**User Autonomy**	**Comprehensive Economic System**	**Personalized Customization**	**Currency System Design**	**Interface layout customization**	**Transaction security**	**Overall Score**
S1	0.6	0.8	0.6	0.8	0.8	0.8	0.6
S2	0.8	0.6	0.4	0.6	0.6	0.8	0.8
S3	0.6	0.2	0.4	0.6	0.6	0.8	1
S4	0.4	0.4	0.4	0.6	0.8	0.6	0.8
...	...	...	...	...	...	...	...
S603	1	1	0.8	0.4	0.6	0.6	0.4
S604	0.4	0.6	0.4	0.6	0.2	0.8	0.6
S605	0.4	0.8	0.4	0.6	0.2	0.8	0.6
S606	0.8	0.8	0.6	0.4	0.8	0.4	0.8

Table 6: CH Index for Different Numbers of Clusters

Number of Clusters	2	3	4	5	6	7	8	9
CH Index	611.81	679.03	697.33	796.63	792.18	775.06	766.09	758.23

Table 7: Number of Samples in Each of the 5 Clusters

Clusters	1	2	3	4	5
Samples	45	56	416	30	50

highest number of samples as our sample set for subsequent experiments. The distribution of samples in each class is shown in Table 7.

Class 3 contains 416 samples, making it the class with the highest number of samples. We will use the samples from this class as our dataset for future experiments. The dataset is divided into training and testing samples, with 80% allocated to training and 20% to testing. The T BP neural network weak learners, constructed based on the dataset, are then trained.

4.3 AdaBoost Algorithm Structure Design

Before applying the metaverse product maturity assessment model, which is based on the AdaBoost algorithm, to the Virtual Reality (VR) social platform A, it is essential to determine the network parameters that are suitable for the model. Considering the specific product context and striving for optimal model performance, the objective is to tailor the solution to the actual requirements of Product A, while maximizing computational efficiency and prediction accuracy.

(1) Determining the Number of Base Learners:

Each BP neural network is considered one of the T base learners. T is selected from the range [5, 10, 15, 20, 25, 30], and after multiple training iterations for each number of base learners within the specified range, it is determined that T=10 achieves the highest prediction accuracy.

(2) Determining BP Neural Network Parameters:

This includes parameters such as the number of layers in the neural network, nodes in each layer, node transfer functions, and training functions.

(3) Determining Nodes in Each Layer of a BP Neural Network:

Input Layer Nodes: The evaluation indicator system constructed in this study includes 70 dimensions of input data, resulting in 70 neurons in the input layer.

Output Layer Nodes: The output of the evaluation model in this study is a single-dimensional overall maturity score of the product. Therefore, the number of neurons in the output layer is 1.

In summary, the network parameters for the metaverse product maturity assessment model based on the AdaBoost algorithm are presented in Table 8.

Table 8: Network Parameters for the Model based on the AdaBoost Algorithm

Base Learners	Network Layers	Input Layer Nodes	Output Layer Nodes	Hidden Layer Nodes	Learning Rate
10	3	70	1	12	0.1

4.4 Analysis of Maturity Prediction Results

We trained the model using the sample data from the training set, and the results obtained are presented in Table 9. In the table, the "user's true rating" represents the overall maturity score assigned by the user to the product. The model's predicted values closely match the users' true ratings during the training phase, with an error of less than 0.1.

Table 9: Actual Output Values and Expected Output Values of the Network

Sample data	Actual Output Values N	User's True Rating E	Error(N-E)
D1	59.930	60	0.070
D2	43.928	44	0.072
D3	56.024	56	0.024
D4	54.920	55	0.080
D5	67.063	67	0.063
D6	56.975	57	0.025
D7	59.930	62	0.070
D8	55.089	54	0.011
D9	61.055	61	0.055
D10	47.927	48	0.073

To assess the model's performance, Figures 3 and 4 illustrate the regression analysis of actual output values versus expected output values, as well as comparison graphs displaying predicted values for the testing samples and their corresponding real values. In both graphs, the vertical axis represents the actual output values of the network, while the horizontal axis represents the expected output values. The parameter R represents the correlation between the measured output and the target, with 1 indicating a strong correlation and 0 indicating a random correlation.

From Figure 4, it can be observed that all four correlation coefficients (R) are greater than 0.99. According to the concept of correlation coefficients, there is a strong correlation between the actual output values and users' true ratings. Upon comparing the predicted and actual values for the test samples in Figure 5, it is evident that the errors between them are minimal, and the numerical distributions are essentially consistent. This indicates that the AdaBoost model shows good network performance and validates its feasibility for assessing the maturity of metaverse products.

Additionally, we compared our model with an approach that utilizes average weights for each indicator. For example, the weight of a secondary indicator is calculated as the number of tertiary indicators under that secondary indicator divided by the total number of tertiary indicators. Figure 6 illustrates the comparison results.

The results indicate that the weights learned by our model, based on user ratings for each indicator, result in overall maturity scores that are more closely aligned with users' maturity ratings for the product. In contrast, the approach using average weights assigns equal weight to each indicator, leading to a greater disparity compared to users' true ratings for product maturity.

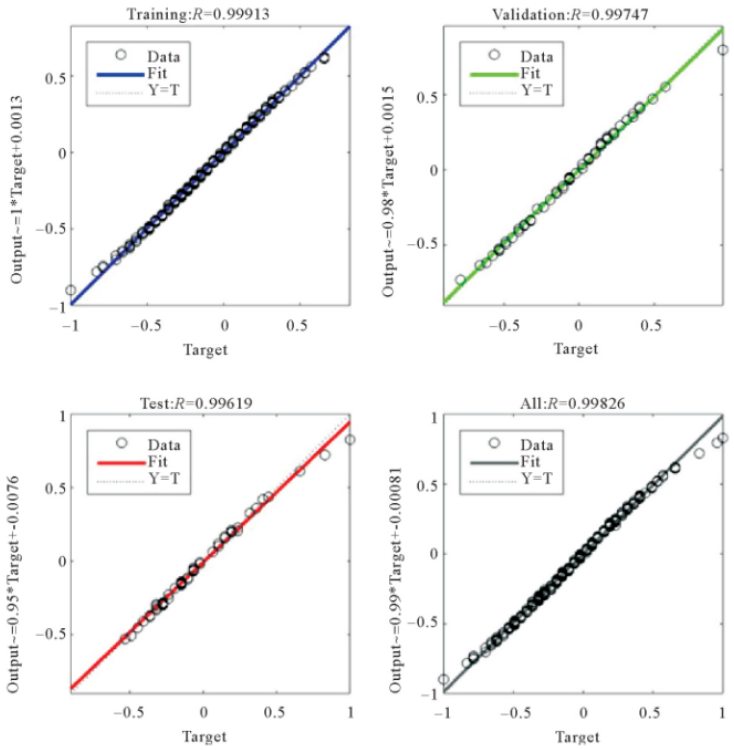

Figure 4: Regression Analysis of Actual Output Values vs. Expected Output Values

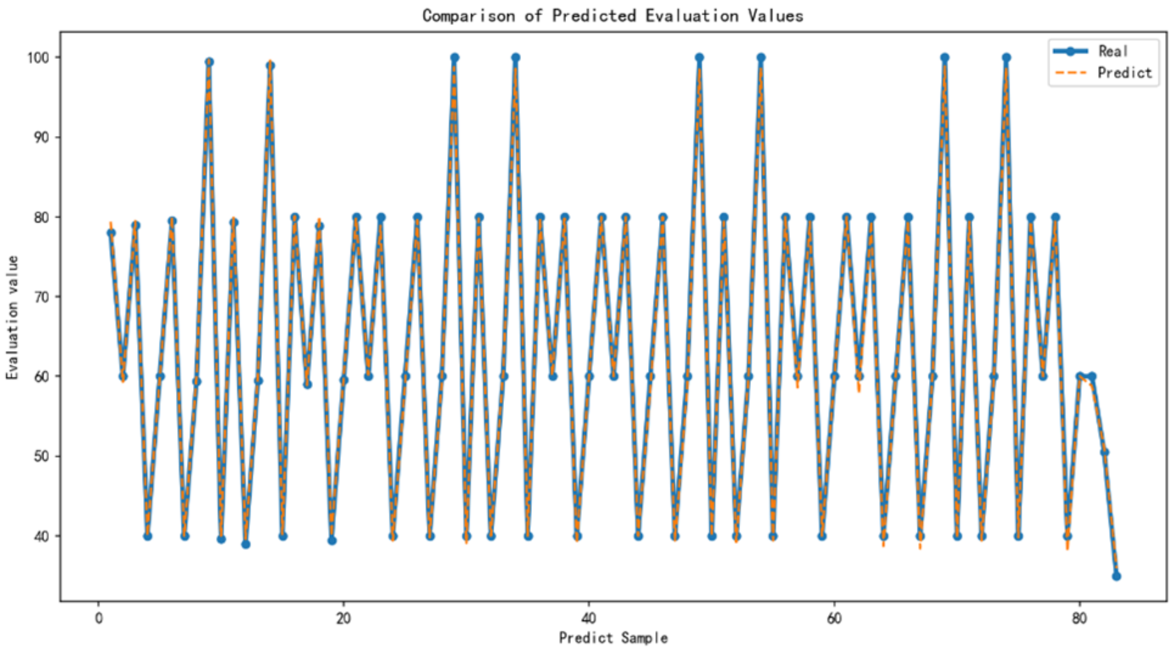

Figure 5: Comparison between Predicted and Actual Values of Samples

5 Conclusions

This paper presents a three-level maturity evaluation system for metaverse products and proposes a maturity evaluation method based on clustering and the AdaBoost algorithm(K-means-AdaBoost). By learning the objective weights of indicators at each level, it is possible to avoid the subjectivity of manual evaluations. In comparison, our approach is more objective and closer to users' true ratings than calculating the overall maturity score using average weights for indicators at each level. This improves the accuracy and efficiency of assessing metaverse product maturity.

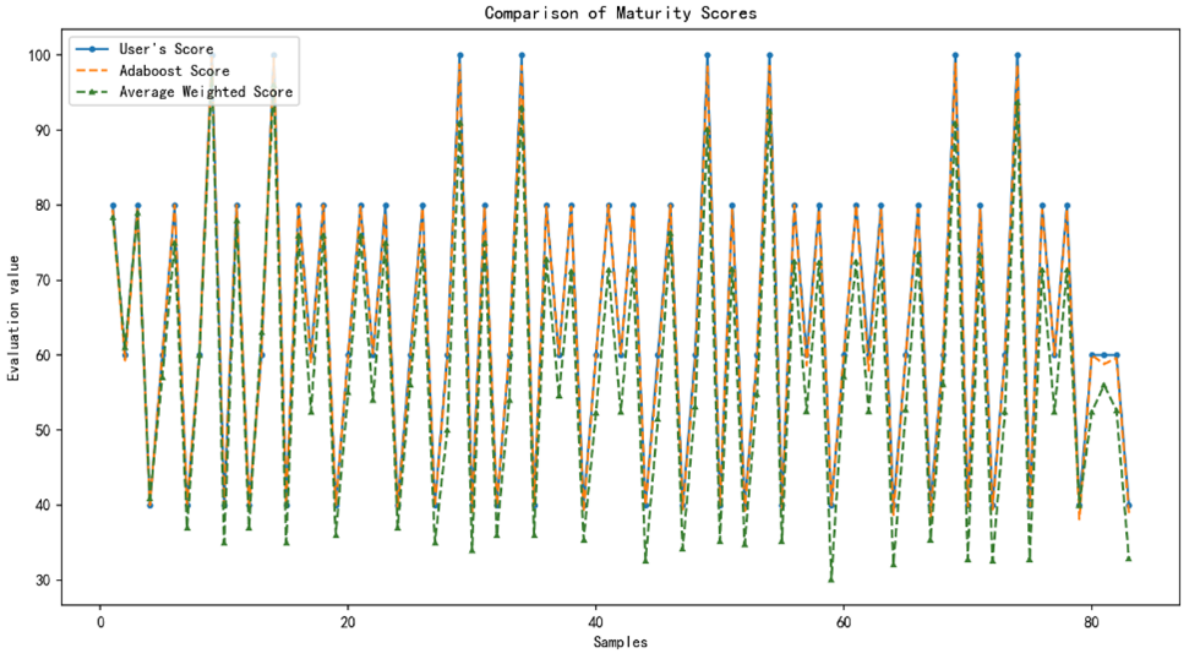

Figure 6: Comparison of Maturity Scores between Our Model and the Average Weighting Method

However, as the complexity of the metaverse continues to grow, data-driven assessment methods face a number of serious challenges, such as data quality, standardization, timeliness, model selection and social subjectivity etc. Data-driven metaverse maturity assessment methods rely on a large amount of data, while data quality and reliability are challenged by multiple sources of data in the metaverse . And the integration of the heterogeneous data is a complex task, which may be affected by different platforms, applications, and user behaviors. Different metaverse platforms adopt different data standards and formats, which leads to difficulties in data integration and analysis.

Due to the rapid development of metaverse technologies, the assessment methodologies should also be continuously updated to better reflect the situation of metaverse system. The outdated assessment methods may not accurately reflect the maturity of the current metaverse, leading to distorted assessment results. Therefore, assessment methods need to be synchronized with technological development to ensure their validity[22]. Currently, the metaverse maturity assessment methodology is still in its infancy, but with a deeper understanding of metaverse data, we believe the assessment methods will evolve to better adapt to the development needs of metaverse systems. Of course, timely access the user feedback and needs will also be a key component of the assessment, making the metaverse system more closely to users' expectations. At the same time, multidisciplinary integration will be a future trend as data-driven maturity assessment will require expertise from big data[23],software testing[24],information fusion[25],fuzzy logic[26],Collaborative decision-making[27],style analysis[28]for a more comprehensive and in-depth understanding[29] of the metaverse system.

Conflict of interest

The authors declare no conflict of interest.

References

[1] Meng, Z.; She, C.; Zhao, G. & Martini, D. (2022). Sampling, Communication, and Prediction Co-Design for Synchronizing the Real-World Device and Digital Model in Metaverse, *IEEE Journal on Selected Areas in Communications*, 41, 288-300.

[2] Star X.Zhao; Qiao Lili; Fred Y.Ye.(2022). A Review of Metaverse Research and Applications, *Journal of Information Resources Management*, 2022, 12(4): 12-23,45.

[3] Wang, Y.; Su Z.; Zhang N. et al.(2022). A Survey on Metaverse: Fundamentals, Security, and Privacy, *IEEE Communications Surveys & Tutorials*, 2022, 25: 319-52.

[4] Lee, L-H; Braud, T.; Zhou, P. et al. (2021). All One Needs to Know about Metaverse: A Complete Survey on Technological Singularity, Virtual Ecosystem, and Research Agenda, *ArXiv*, 2021, abs/2110.05352.

[5] Wang, H.; Ning, H.; Lin, Y. et al. (2023). A Survey on the Metaverse: The State-of-the-Art, Technologies, Applications, and Challenges, *IEEE Internet of Things Journal*, 2023, 10: 14671-88.

[6] Dionisio, J. D. N.; III, W. G. B., Gilbert, R. (2013). 3D Virtual worlds and the metaverse: Current status and future possibilities, *ACM Comput Surv*, 2013, 45(3): Article 34.

[7] Park, S-M; Kim, Y-G. (2022). A Metaverse: Taxonomy, Components, Applications, and Open Challenges, *IEEE Access*, 2022, 10: 4209-51.

[8] Zainab, H.E.; Bawany, N.Z.; Imran, J. & Rehman, W. (2022). Virtual Dimension—A Primer to Metaverse, *IT Professional*, 24, 27-33.

[9] Weinberger, M.; Gross, D. (2023). A Metaverse Maturity Model, *Global Journal of Computer Science and Technology*, 2023.

[10] Pamučar, D.; Deveci, M.; Gokasar, I. et al. (2022). A metaverse assessment model for sustainable transportation using ordinal priority approach and Aczel-Alsina norms, *Technological Forecasting and Social Change*, 2022.

[11] Jin, Y.; Zhang, H.; Zhao, L. et al. (2023). Data driven improvement of user perception in mobile business halls, *Communication World*, 2023, (20): 40-1.

[12] Song, Y.N.; Li, Z.; Wang, Y. et al. (2023). Research on data-driven spatial situational entity cognition methods, *The 11th China Command and Control Conference, Beijing, China*, F, 2023 [C].

[13] Li, Y.P. (2023). Data driven decision-making: key factors in smart library management, *Culture Monthly*, 2023, (09): 120-2.

[14] Wu, H.; Liu, J.H.; Zhu, H. et al. (2023). Data driven traffic signal control perception evaluation diagnosis optimization closed-loop technology (II): State perception, *Road Traffic Management*, 2023, (09): 28-31.

[15] Yu, J. E. (2022). Exploration of educational possibilities by four metaverse types in physical education, *Technologies*, 10(5), 104.

[16] VİSCONTİ, R. M. (2022). From physical reality to the Metaverse: a Multilayer Network Valuation, *Journal of Metaverse*, 2(1), 16-22.

[17] Eom, H.; Kim, K.; Lee, S. et al. (2019). Development of virtual reality continuous performance test utilizing social cues for children and adolescents with attention-deficit/hyperactivity disorder, *Cyberpsychology, Behavior, and Social Networking*, 2019, 22(3): 198-204.

[18] SOLOMATINE, D. P.; SEE, L. M.; ABRAHART, R. J. (2017). Chapter 2 Data-Driven Modelling, *Concepts , Approaches and Experiences*, F, 2017 [C].

[19] Park, S. M. & Kim, Y. G. (2022). A Metaverse: Taxonomy, Components, Applications, and Open Challenges, *IEEE Access*, 10, 4209-4251. https://doi.org/10.1109/ACCESS.2021.3140175

[20] DAMAR, M. (2021). Metaverse Shape of Your Life for Future: A bibliometric snapshot, *Journal of Metaverse*, 1(1), 1-8.

[21] Njoku, J. N.; Nwakanma, C. I.; Amaizu, G. C. & Kim, D. S. (2023). Prospects and challenges of Metaverse application in data-driven intelligent transportation systems, *IET Intelligent Transport Systems*, 17(1), 1-21.

[22] Guston, D. H. & Sarewitz, D. (2020). Real-time technology assessment, In *Emerging Technologies*, (pp. 231-247). Routledge.

[23] Zhang, D. (2017). High-speed train control system big data analysis based on the fuzzy RDF model and uncertain reasoning, *International Journal of Computers Communications & Control*, 12(4), 577-591.

[24] Zhang, D.; Sui, J. & Gong, Y. (2017). Large scale software test data generation based on collective constraint and weighted combination method, *Tehnicki Vjesnik/Technical Gazette*, 24(4).

[25] Chen, W. (2023). Deep adversarial neural network model based on information fusion for music sentiment analysis, *Computer Science and Information Systems*, (00), 31-31.

[26] Dzitac, I.; Filip, F. G. & Manolescu, M. J. (2017). Fuzzy logic is not fuzzy: World-renowned computer scientist Lotfi A. Zadeh, *International Journal of Computers Communications & Control*, 12(6), 748-789.

[27] Filip, F. G. (2022). Collaborative decision-making: concepts and supporting information and communication technology tools and systems, *International Journal of Computers Communications & Control*, 17(2).

[28] Constantinescu, Z.; Marinoiu, C. & Vladoiu, M. (2010). Driving style analysis using data mining techniques, *International Journal of Computers Communications & Control*, 5(5), 654-663.

[29] Lyu, Z. (2023). State-of-the-art human-computer-interaction in metaverse, *International Journal of Human–Computer Interaction*, 1-19.

Cite this paper as:

Tang, M.C.; Cao, J.; Fan, Z.X.; Zhang, D.L.; Pandelica, I. (2024).A Data-Driven Assessment Model for Metaverse Maturity, *International Journal of Computers Communications & Control*, 19(4), 6498, 2024.

https://doi.org/10.15837/ijccc.2024.4.6498