www.cad-journal.net
Computer-Aided Design AND APPLICATIONS

A Deep Learning Algorithm for Music Information Retrieval Recommendation System

Hongxia Liu[1] and Chenjing Zhao[2]

[1]School of music and dance, Xinyang Normal University, Henan 464000, China, lhx@xynu.edu.cn
[2]School of music and dance, Xinyang Normal University, Henan 464000, China, z851447528@163.com

Corresponding author: Hongxia Liu, lhx@xynu.edu.cn

Abstract. In music information retrieval (MIR), deep learning (DL) can use its powerful feature extraction ability to extract more effective features from music data. The traditional MIR method can't effectively deal with the diversity, dynamics, heterogeneity and real-time characteristics of Internet of Things (IoT) data. This article proposes a DL-based MIR and IoT recommendation system, aiming at providing more accurate and efficient computer-aided MIR and recommendation services. In order to evaluate the performance of IoT music recommendation system based on different recommendation algorithms, the advantages and disadvantages of various recommendation algorithms are analyzed through comparative experiments and system load test experiments. The results show that DL algorithm has obvious advantages in all indicators, followed by Collaborative Filtering (CF) algorithm and matrix decomposition-based algorithm, while the recommendation effect of K-nearest neighbor algorithm is relatively poor. The implementation of this algorithm can provide useful reference and support for the design and optimization of IoT recommendation system, improve the performance and stability of recommendation system, and thus promote the sustainable growth of music industry.

Keywords: Music Information Retrieval; Internet of Things; Recommendation System; Deep Learning; Computer Aided
DOI: https://doi.org/10.14733/cadaps.2024.S13.1-16

1 INTRODUCTION

The Internet of Things (IoT), also known as the third wave of the world information industry, is another significant progress in the growth of the world information industry after computers and the Internet. It connects countless real objects together and forms a super-large network, which provides people with unprecedented information acquisition and analysis and processing capabilities. Supervised machine learning is a machine learning method that analyzes a large amount of labeled data for model training and utilizes the trained model for classification or prediction. In the Internet

© 2024 U-turn Press LLC, http://www.cad-journal.net

of Things, supervised machine learning can be used to detect abnormal behavior, identify malicious software patterns, and predict potential internal attacks. A common method to alleviate malicious internal attacks in the Internet of Things is to establish an anomaly detection model. These models are trained to recognize normal and abnormal behavior, and when abnormal behavior is detected, the system triggers an alarm. For example, using the Support Vector Machine (SVM) algorithm in supervised learning, a classifier can be trained to distinguish between normal device behavior and abnormal behavior. In addition, supervised learning algorithms such as random forest and gradient enhancement can also be used for anomaly detection. Ahmad and Shah [1] establish a predictive model to predict potential internal attacks. This model can be based on time series analysis or regression algorithms to predict future device behavior. For example, by analyzing the device's battery usage or network traffic data, it can be predicted whether the device is about to be attacked by malicious software. With the rapid growth of science and technology, human society is entering an IoT era, and everything is interconnected, and information acquisition and exchange become more convenient. With the popularization of cloud computing, more and more enterprises and individuals are moving their business and data to the cloud. However, this trend also brings new security challenges, with Economic Denial of Sustainability Attacks (EDSA) being an increasingly serious threat. EDSA attacks aim to achieve attackers' goals by disrupting or disrupting the economic sustainability of cloud computing environments, such as reducing service levels and increasing latency. Therefore, developing a system that can effectively detect and defend against EDSA attacks is particularly important. Aldhyani and Alkahtani [2] explore an economic denial of sustainability attack detection system based on artificial intelligence algorithms, and provide a detailed introduction to its application in cloud computing environments. At present, the detection methods for EDSA attacks mainly focus on anomaly detection, traffic analysis, model monitoring, and other fields. The anomaly detection method detects abnormal behavior that does not match normal behavior by monitoring the normal operation status of the system. IoT is widely used in industry, agriculture, medical care, transportation and other fields by connecting various intelligent sensing devices through the Internet and mobile communication networks to realize the functions of remote information transmission and control of articles. IoT not only facilitates people's life, but also brings new features such as data diversity, dynamics, heterogeneity and real-time, which brings new challenges to information retrieval technology.

Sound is an indispensable part of human life and plays an important role in people's mood, mood and entertainment. With the growth of network technology, the number and variety of music resources have doubled, making it more difficult for people to find the music they need. Machine learning algorithms automatically learn and improve models by analyzing a large amount of data, without the need for human feature engineering. Deep learning algorithms utilize neural network models to high-level abstract data, thereby better identifying complex network attacks. In the field of network security, the application of machine learning and deep learning algorithms mainly includes intrusion detection, malware analysis, vulnerability mining, and other aspects. The implementation of the system is based on the Python language and TensorFlow deep learning framework. Firstly, the data collection function is implemented through web crawler technology. Then, use the Pandas library for data preprocessing, such as data cleaning, label encoding, etc. Finally, a deep neural network model was used for training and testing. In the experimental environment, Alkahtani and Aldhyani [3] tested the system and the results showed that the system was able to accurately identify most network attacks. For certain complex attack behaviors, the system also has strong robustness and generalization ability. Through comparative experiments, we found that deep learning algorithms have higher accuracy and efficiency in network security detection compared to traditional machine learning algorithms. MIR uses computer science, digital signal processing, artificial intelligence and other related technologies to extract useful information from a large quantity of music resources to help users find the required content quickly and accurately. With the continuous progress of science and technology, the application of the Internet of Things and edge computing in the field of health care has attracted more and more attention. By combining IoT devices and edge computing technology, real-time monitoring, analysis and processing of medical data can be realized, thus providing more accurate and efficient diagnosis and treatment solutions for medical staff. Alnaim and

© 2024 U-turn Press LLC, http://www.cad-journal.net

Alwakeel [4] introduced the application field and significance of the Internet of Things based on machine learning - edge computing healthcare solution. In the application of the Internet of Things, machine learning algorithms are widely used in data mining, pattern recognition, and predictive analysis. In the environment of edge computing, machine learning algorithms can better meet the high requirements of data processing speed and accuracy in the field of health care. Deep learning algorithms such as convolutional neural networks (CNN) and recurrent neural networks (RNN) can be used in fields such as medical imaging and speech recognition.

However, the traditional MIR method can not effectively deal with the diversity, dynamics, heterogeneity and real-time characteristics of IoT data, so it is difficult to meet the needs of the IoT era. With the rapid development and widespread application of Internet of Things technology, network attacks are also showing an increasing trend. However, in recent years, the rise of machine learning technology is gradually changing the nature of IoT network attacks. Bout et al. [5] explored the application and classification of machine learning in IoT network attacks, and analyzed how machine learning can alter the properties of IoT network attacks. Machine learning is an advanced technology that can automatically extract knowledge or patterns from a large amount of data. Machine learning algorithms can automatically detect abnormal behavior, such as sudden traffic surges or abnormal device connections, in order to alert potential network attacks. By analyzing historical data, machine learning algorithms can learn normal behavior patterns and automatically identify behaviors that do not match these patterns. Through machine learning algorithms, risk assessment can be conducted on IoT devices to determine which devices may be vulnerable to attacks. In addition, machine learning can also predict emerging network attack methods and provide preventive measures. When machine learning algorithms detect potential network attacks, they can automatically trigger defense mechanisms, such as shutting down the attacked device or reconfiguring the network to prevent the attack. This article proposes a DL-based MIR and IoT recommendation system, which aims to solve the challenges brought by the diversity, dynamics, heterogeneity and real-time performance of IoT data and provide more accurate and efficient computer-aided MIR and recommendation services.

DL is a branch of machine learning, which uses artificial neural network to simulate the operation mode of human brain neural network and realize the processing and analysis of complex data. The concept of complexity matching refers to the matching relationship between the complexity of information processing and the complexity and dynamics of the environment in the biological brain. Carpentier et al. [6] explored how brain signals reflect this complex matching during music listening and reward processes. In the process of understanding music, the brain needs to process complex information flows from the auditory system. Music encompasses various elements, such as melody, rhythm, harmony, and timbre, which interact in complex ways to form what we call 'music'. The brain's processing of this information is not singular or linear, but involves complex neural network activities, including the collaborative work of multiple regions such as the auditory cortex, frontal lobe, and temporal lobe. Interestingly, this complexity processing of music does not only occur in the auditory system. When we enjoy music, the brain's reward system (mainly including the midbrain limbic dopamine system) is also activated. This system is related to our core functions of emotional, motivational, and behavioral control, and its activities provide us with a sense of pleasure and satisfaction. In MIR, DL can use its powerful feature extraction ability to extract more effective features from music data and improve the retrieval accuracy. For example, by training deep neural network (DNN) to learn the inherent laws and patterns of music features, a more accurate MIR engine can be established. Dake [7] explored the capabilities and advantages of 5G networks supporting IoT devices running in heterogeneous network systems. By comparing the characteristics of traditional communication technologies and 5G networks, we found that 5G networks have higher data transmission rates, lower latency, and larger network capacity. This enables 5G networks to support more IoT devices connected simultaneously and provide more real-time and efficient data transmission services. However, 5G networks also have some shortcomings, such as relatively small coverage and high device power consumption. In order to fully leverage the advantages of 5G networks supporting IoT devices running in heterogeneous network systems, it is necessary to increase the construction of 5G networks, optimize the power consumption design of devices, and

© 2024 U-turn Press LLC, http://www.cad-journal.net

improve the energy efficiency ratio of devices. In the aspect of music recommendation system, DL can realize more personalized music recommendation service by analyzing users' interests, behaviors and habits, as well as the content, style and interpretation of music works. For example, using the depth CF algorithm to analyze the similarity between users and music works can provide users with music recommendations that are more suitable for their needs.

Dang et al.'s [8] cloud based digital pairing for structural health monitoring using deep learning has brought new breakthroughs in the field of structural health monitoring, achieving more efficient and intelligent monitoring. By combining deep learning with cloud computing technology, it is possible to improve the accuracy and efficiency of monitoring, reduce costs, and achieve automated data processing and analysis. In the future, with the continuous development of deep learning and cloud computing technology, cloud based digital pairing using deep learning for structural health monitoring is expected to be widely applied in more fields. Bringing greater convenience and security to the development of society and the lives of the people. In large-scale mechanical equipment, monitoring and analyzing its operating and vibration data can effectively predict equipment failures and maintenance needs. In addition, deep learning combined with cloud computing can also be used to achieve structural health monitoring and early warning in fields such as building structures and petrochemicals. The purpose of this article is to study a computer-aided MIR and IoT recommendation system based on DL. Firstly, how to use DL algorithm to process and extract music data features in IoT is studied. Then, DL is applied to MIR and recommendation system to improve the accuracy and efficiency of retrieval and recommendation. Finally, the effectiveness and feasibility of the proposed method are verified by experiments. The innovations of this study mainly include the following aspects:

(1) This article proposes a method of recommendation system for MIR and IoT based on DL. Through DL algorithm, this method not only realizes the feature extraction and processing of music data, but also realizes more accurate and personalized computer-aided MIR and recommendation.

(2) In the aspect of music feature extraction, this article proposes a new DL model, which can better capture the time domain and frequency domain features of music data, thus improving the accuracy and efficiency of music feature extraction.

(3) The effectiveness and feasibility of the proposed method are verified by experiments. The experimental results show that the method proposed in this article has high accuracy and efficiency in both MIR and recommendation, which provides new ideas and methods for the growth of MIR and recommendation technology in IoT era.

The structure of the article is as follows: The first section is the introduction, introducing the related background of IoT and MIR; The second section is a summary of related research, which summarizes the application research of DL in MIR and recommendation system in recent years. The third section is music data feature extraction and processing based on DL. The fourth section is based on computer-aided MIR and recommendation algorithm; The fifth section is experiment and analysis; The sixth section is the conclusion and prospect.

2 OVERVIEW OF RELEVANT RESEARCH

In recent years, the application research of DL in MIR and recommendation system has made remarkable progress. DL algorithm can automatically learn and extract the intrinsic features of music data, thus improving the accuracy of MIR and recommendation. In addition, DL can also handle unstructured music data, such as audio and lyrics, providing more possibilities for MIR and recommendation.

2.1 Application of DL in MIR

Traditional music feature extraction methods usually need to manually define feature extraction rules, and it is difficult to capture the complex features of music data. DL can automatically learn and extract the features of music data by training neural networks, thus avoiding the tedious process of manually defining features. Gyamfi and Jurcut [9] introduce the intrusion detection in the Internet of Things system, using the design methods of multi access edge computing, machine learning and data sets. In traditional IoT systems, intrusion detection typically relies on central servers for centralized processing. However, with the expansion of the scale of the Internet of Things, this centralized processing method faces problems of low efficiency and high latency. Multi access edge computing (MEC) is a new computing mode, which allocates computing tasks to edge devices for processing, thus reducing the burden of the central server. In IoT systems, utilizing MEC technology can improve the efficiency and response speed of intrusion detection. Machine learning is a data-driven technology that automatically discovers patterns and patterns by analyzing large amounts of data. In IoT intrusion detection, machine learning algorithms can automatically identify abnormal behavior and detect potential intrusions. Hernandez et al. [10] compared the application effects of CNN, LSTM, and BiLSTM deep learning methods in tone analysis in polyphonic automatic music transcription. The results indicate that all three methods have certain advantages, but there are also some limitations. Specifically, CNN has high accuracy and stability in extracting audio signal features, but its temporal feature extraction ability is weak. LSTM is good at processing temporal features of audio signals, but it is not accurate enough in static feature extraction. BiLSTM combines the advantages of CNN and LSTM, but training is difficult and requires more training data and computing resources. The future application directions of deep learning in tone analysis in polyphonic automatic music transcription may include: 1) combining multiple deep learning models to form complementary advantages and improve the accuracy and comprehensiveness of tone analysis. 2) Conduct in-depth research on the internal mechanisms of deep learning models, optimize model structure, algorithms, and training methods, and improve the learning and generalization abilities of the models.

DL can deal with the high-dimensional features in music data and automatically learn the music similarity calculation model. These models can effectively measure the similarity between music data, thus supporting the MIR task. Music object recognition refers to the automatic recognition of various elements in music through computer algorithms, such as notes, chords, timbre, rhythm, etc. The rise of deep learning technology provides new solutions for music object recognition. By establishing multi-layer neural networks, deep learning models can automatically learn music features and achieve accurate recognition of music objects. Huang et al. [11] transformed the original audio data into a form suitable for model input. Usually involves operations such as audio signal preprocessing, feature extraction, and dimension reduction. For example, techniques such as Fourier transform or Mel frequency cepstrum coefficient (MFCC) are used to extract features such as melody, rhythm, and timbre of audio. At this stage, the preprocessed data is input into the deep learning model for training. These models can automatically learn music features and optimize and adjust them based on training data. The trained model can be used to predict new music data. According to application requirements, the model can output different types of music object recognition results, such as song classification, singer recognition, and music style judgment. Machine learning is an artificial intelligence method that automatically learns patterns in data and makes predictions or classifications through training models. Machine learning has applications in many fields, such as computer vision, natural language processing, and medical diagnosis. However, machine learning faces some challenges in practical applications, such as data quality and diversity, computational resources, and algorithm optimization. In addition, the black box nature of machine learning also limits its application in certain fields, such as finance and healthcare, which require interpretive decision-making processes, which are precisely what machine learning models lack. Software Defined Network (SDN) is a network architecture that separates the network control plane from the data plane, making network configuration and management more flexible and programmable. The characteristics of SDN make it widely applicable in scenarios such as cloud computing, data centers, and wide area networks. However, SDN also faces some challenges, including security and privacy protection, the impact of network performance and stability, as well as standardization and compatibility issues. In addition, the deployment and management of SDN also require more automation and intelligent means to meet the needs of large-scale networks [12].

DL can be used in the task of music semantic understanding, that is, extracting and identifying information with semantic meaning from music data. Lozano et al. [13] introduced the application

© 2024 U-turn Press LLC, http://www.cad-journal.net

and significance of a context aware recommendation system in the music field. Context aware recommendation system is an intelligent system that recommends based on user historical behavior, real-time contextual information, and machine learning algorithms. In the field of music, the system can understand users' music preferences, listening history, playlists, and other information, and develop personalized music recommendations based on these data. Feature extraction and selection are key factors that affect the performance of music recommendation systems. These methods can extract effective features from raw data, reflecting users' music preferences and behavioral patterns. Model training is the core step in building a recommendation system. Context aware recommendation systems typically use deep learning models for training. During the training process, annotated data is used as input to minimize prediction errors (such as mean square error) through optimization algorithms such as random gradient descent. And designed and evaluated a music recommendation system using self-encoder and generative countermeasure network (GAN).

2.2 Application of DL in Music Recommendation System

DL can automatically learn and model users' musical interest preferences by analyzing users' historical behavior data and other related information. In this situation, deep learning, as a powerful machine learning technology, provides new ideas for solving this problem. Mao et al. [14] proposed a joint channel estimation and active user detection method based on deep learning. Firstly, this method transforms the channel estimation problem into a regression problem and establishes a multi-layer neural network (NN) model for channel estimation. This model can learn and simulate the actual channel environment, thereby obtaining more accurate CSI. In addition, by using deep learning techniques, this method can adaptively handle environmental noise and other interference factors, thereby improving the accuracy of channel estimation. Secondly, for the user detection problem, this method adopts a binary NN model. This model can distinguish between active and inactive users, thereby achieving active detection of users. In order to improve the classification accuracy of the model, this method also introduces the concept of Long Term Dependency (LTR), enabling the model to better understand and utilize the time series information of user behavior. DL can be used to implement multi-task recommendation, such as multi-task self-encoder, and cooperative processing of multiple recommendation tasks can be achieved through joint training of multiple self-encoders. With the popularization and in-depth application of the Internet of Things (IoT), the challenges to data processing are also increasing. Traditional data processing methods cannot effectively handle such large-scale and complex data. In recent years, the rapid development of Adaptive Deep Neural Networks (ADNN) has provided an effective way to solve this problem. Pandit et al. [15] explored the application advantages and feasibility of adaptive deep neural networks in the field of the Internet of Things. Adaptive deep neural network is a deep learning model that can automatically adjust its structure and parameters based on input data. Compared with traditional deep neural networks, adaptive deep neural networks have better flexibility and adaptability. It can automatically adjust the network structure and optimize parameter settings based on different data features to improve prediction accuracy and processing efficiency. Adaptive deep neural networks can accurately predict the failure time and type of equipment by automatically learning the operating status data of the equipment, thereby carrying out maintenance and replacement in advance and avoiding production interruptions.

With the development and popularization of network technology, computer-aided teaching management systems have been widely applied in the field of education. Especially in music appreciation courses, network resources are abundant, and computer-aided teaching management systems can help teachers better manage course resources, improve teaching quality, and also help students better appreciate and understand music. Pei and Wang [16] analyzed the computer-assisted teaching management system for music appreciation courses based on network resources and proposed corresponding design suggestions. The system needs to provide functions for uploading, downloading, classifying, and searching music resources, so that teachers and students can easily access the required music resources. The system needs to support multiple music playback formats and provide high-quality music playback and appreciation experiences. Teachers can develop teaching plans through the system, arrange course content and teaching progress, and

© 2024 U-turn Press LLC, http://www.cad-journal.net

supervise and manage students' learning situation. The combination of intelligent city IoT and deep learning algorithms has achieved intelligent management and optimization in various fields of the city. For example, in the field of intelligent transportation, using deep learning algorithms to mine and analyze traffic data can achieve functions such as intelligent traffic signal control and intelligent vehicle scheduling; In the field of intelligent healthcare, medical data is collected through intelligent IoT devices, and deep learning algorithms are used for medical image analysis, disease prediction, and more. However, the application of smart city IoT and deep learning algorithms also faces some challenges. Firstly, data security and privacy protection issues are prominent. A large amount of urban data involves personal privacy and trade secrets, and ensuring data security and privacy protection is a major challenge. Secondly, the technological maturity needs to be improved. Although smart city IoT and deep learning algorithms have been successfully applied in many fields, they still need to be improved in terms of stability, reliability, and practicality. Finally, urban planning and management need to be more intelligent and refined. The application of smart city IoT and deep learning algorithms will generate a large amount of data. How to effectively process and analyze this data to provide intelligent support for urban planning and management work is also an important challenge [17].

3 DL BASED FEATURE EXTRACTION AND PROCESSING OF MUSIC DATA

With the continuous growth of music resources, people need more intelligent ways to obtain suitable music resources and enhance music experience. MIR technology provides better intelligent services for music practitioners, music platforms and fans. Music data has diversity and complexity, and traditional feature extraction methods are often difficult to effectively capture its internal features. The application of deep learning in tone analysis in polyphonic automatic music transcription still faces some challenges. The preprocessing and feature extraction techniques of audio signals need to be further improved and improved to meet the transcription needs of polyphonic music in different music styles, instruments, and environments. Deep learning models require a large amount of annotated data for training, while there is relatively little annotated data in the field of music transcription, and effective data augmentation and transfer learning methods need to be explored. The interpretability of deep learning models is insufficient, making it difficult to understand and trust their decision-making process. Therefore, it is necessary to strengthen interpretability research [18]. Sun et al. [19] introduced the partition based mobile group sensing task allocation technology for the maintenance of the solar powered pest control lamp IoT. By dividing the monitoring area into different sub areas and utilizing sensor nodes to collaborate, distributed monitoring and data collection of environmental parameters can be achieved. At the same time, through reasonable task allocation strategies and implementation methods, the efficiency and accuracy of task allocation can be effectively improved, and the amount of data transmission and processing can be reduced. Although this technology has achieved certain results and applications, there are still some shortcomings and areas for improvement. Future research directions and issues include improving task allocation strategies, improving task completion efficiency, and enhancing sensor node stability. In recent years, DL has shown great potential in music feature extraction. Compared with the traditional MIR method, DL technology can better handle complex music features and improve the retrieval accuracy and efficiency. When DL is applied to the music recommendation system, it can predict the music that users may be interested in in the future according to their historical records and behaviors, and provide users with more personalized recommendation services.

Energy network refers to a network system that organically integrates energy production, transmission, distribution, consumption, and other links. In energy networks, load forecasting is a very important task, which has a crucial impact on the balance of energy supply and demand, safe operation, and efficiency improvement. Traditional load forecasting methods are usually based on statistics and experience, making it difficult to adapt to complex and ever-changing energy network environments. Zamee et al. [20] studied a data-driven load forecasting method to improve prediction accuracy and stability, which is of great significance for ensuring the stable operation of energy networks and energy conservation and emission reduction. Multi correlation STDNN (MD STDNN) is a

© 2024 U-turn Press LLC, http://www.cad-journal.net

new neural network model that combines multi correlation technology with STDNN, and its application in load forecasting has also achieved good results. This section will introduce the methods of music data feature extraction and processing based on DL in detail, including data preprocessing, feature extraction and training and application of DL model. Before the feature extraction of music data, it is necessary to preprocess the music data, including data cleaning, standardization and framing. Data cleaning is mainly to remove invalid and noisy data from music data to ensure the quality and reliability of data. Standardization is to normalize music data to eliminate the influence of different dimensions on feature extraction. Framing is to divide music data into several small frames according to time series to capture the dynamic characteristics of music data.

Deep belief network (DBN) is a DL model, which is composed of several restricted Boltzmann machines (RBM), and can be used to extract and classify music data. This article uses DBN to extract features from music data, and applies it to MIR and recommendation system. DBN, which is composed of multiple RBM stacks, takes feature vectors as input, and extracts hierarchical features of music data through unsupervised learning step by step. After the training is completed, the hidden layer feature vectors of DBN can be obtained, which can be used as the representation vectors of music data. The DL structure of the constructed music recommendation system is shown in Figure 1.

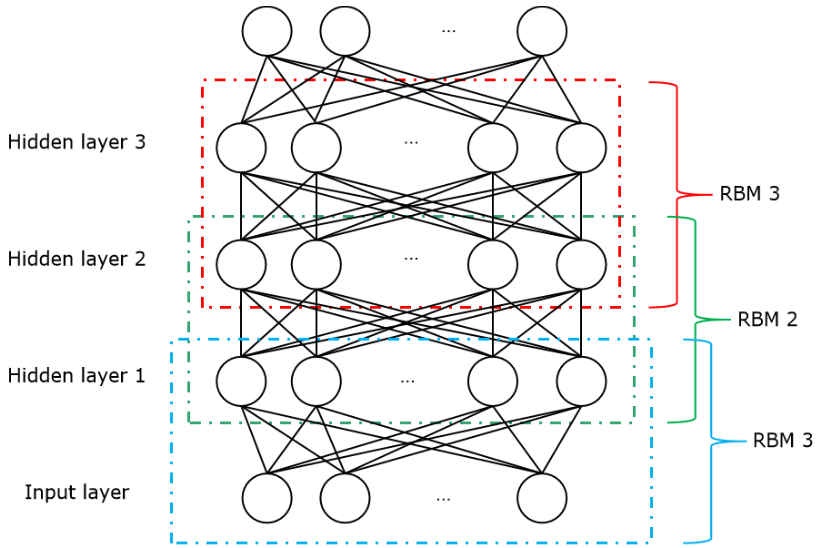

Figure 1: DL structure of music recommendation system.

The overall framework of the recommendation algorithm is shown in Figure 2. The model mainly consists of three stages: structural knowledge feature extraction, joint learning and recommendation list generation. (1) In the structural knowledge feature extraction stage, the entities related to the project are extracted with the user-project historical interaction and Freebase knowledge map in the recommendation system as inputs to construct a sub-knowledge map, and an improved knowledge map embedding method is used to find the potential vector representation of structured knowledge from the sub-graph, so as to obtain the knowledge-aware project representation; (2) In the joint learning stage, the structured embedding vector and the unstructured project feature vector in the knowledge map are integrated together by adding the element positions to construct the final potential vector representation of the project; (3) In the stage of recommendation list generation, the inner product of the final potential vector representation of users and projects is used as the preference probability value of users, and a personalized recommendation list is generated for the target users.

© 2024 U-turn Press LLC, http://www.cad-journal.net

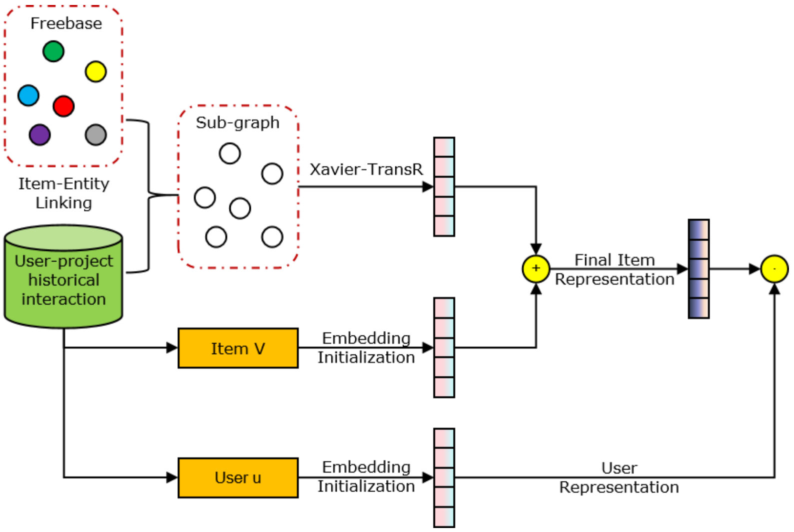

Figure 2: Overall framework of recommendation algorithm.

Firstly, DNN model is used to learn the historical data of users' listening to songs, so as to predict users' future listening behavior, that is, user interest modeling. Then, the CF-based method is used to recommend music works similar to users' interests. In this process, the music representation vector extracted by DBN is used as the feature of music works, and recommendations are made by comparing the similarity between users' interests and music works. Multivariate Gaussian model describes the timbre characteristics of audio signals. In this article, mel-frequency cepstral coefficients (MFCC) is used to calculate the parameters of multivariate Gaussian model. The multivariate Gaussian model parameters of an audio file can be expressed as:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} m_i \tag{1}$$

$$\hat{\Sigma} = \frac{1}{n}\sum_{i=1}^{n} (m_i - \hat{\mu})(m_i - \hat{\mu})^T \tag{2}$$

Where m_i is the MFCC vector and n is the vector dimension. For two n-dimensional Gaussian models $X_1 \sim N(\hat{\mu}_1, \hat{\Sigma}_1)$ and $X_1 \sim N(\hat{\mu}_1, \hat{\Sigma}_1)$, the square root of symmetric KL divergence is used to represent the Gaussian distance d_gl:

$$d\_gl(X_1, X_2) = \sqrt{SKL(X_1|X_2)} \tag{3}$$

$$SKL(X_1|X_2) = \frac{1}{2}KL(X_1|X_2) + \frac{1}{2}KL(X_2|X_1) \tag{4}$$

$$KL(X_1|X_2) = \frac{1}{2}\left(tr(\Sigma_2^{-1}\Sigma_1) + (\mu_2 - \mu_1)^T \Sigma_2^{-1}(\mu_2 - \mu_1) - \log\frac{|\Sigma_1|}{|\Sigma_2|} - n \right) \tag{5}$$

For two FP matrices FP_1 and FP_2, this article uses Euclidean distance to represent FP distance d_fp:

© 2024 U-turn Press LLC, http://www.cad-journal.net

$$d\_fp(FP_1, FP_2) = \sqrt{\sum_i \sum_j \left(fp_1(i,j) - fp_2(i,j) \right)^2} \tag{6}$$

The similar distance d of two audio files is obtained by linear combination:

$$d = 0.7z(d\_gl(X_1, X_2)) + 0.3z(d\_fp(FP_1, FP_2)) \tag{7}$$

Where $z()$ stands for distance normalization function.

4 COMPUTER-AIDED MIR AND RECOMMENDATION ALGORITHM

Music recommendation system is a common application of computer-aided MIR. It predicts users' musical interest preferences by analyzing their listening history, browsing records and other behavioral data, and then recommends music data with similar interests. Through the combination of computer-aided technology and recommendation system, we can process and analyze music data more intelligently, improve the accuracy and efficiency of retrieval and recommendation, and provide people with more personalized music services. According to users' historical records and behaviors, computer-aided technology can predict users' music preferences, recommend more suitable music works for them, and improve the accuracy and efficiency of recommendation. Through computer-aided technology, the music social platform can better analyze the music interests and hobbies among users, recommend more suitable music partners for them, and promote the communication and interaction between users. The combination of computer-aided MIR and recommendation system with IoT can also bring new business opportunities and development space for the music industry.

Computer-aided MIR and recommendation algorithm realize fast and accurate retrieval and recommendation of music data by using computer-aided technology, thus improving the efficiency and accuracy of MIR and recommendation. By analyzing a large quantity of music data, the computer can learn the inherent laws and characteristics of music and automatically generate new music data. The generated new music data can be brand-new songs or fragments of songs. This method needs to train high-quality DL model and high-precision parameter setting to improve the quality and diversity of generated music data.

Users' feedback on projects can generally be divided into two different types: explicit feedback and implicit feedback. Explicit feedback, such as rating, is a strong signal that can represent users' interest preferences and requires users to actively show their interests; However, in many cases, users are generally unwilling to actively grade projects, so it is usually difficult to obtain a large quantity of explicit scoring data. In contrast, implicit feedback such as buying or clicking is more common and easier to obtain in real applications. Explicit feedback and implicit feedback between users and projects are shown in Figure 3.

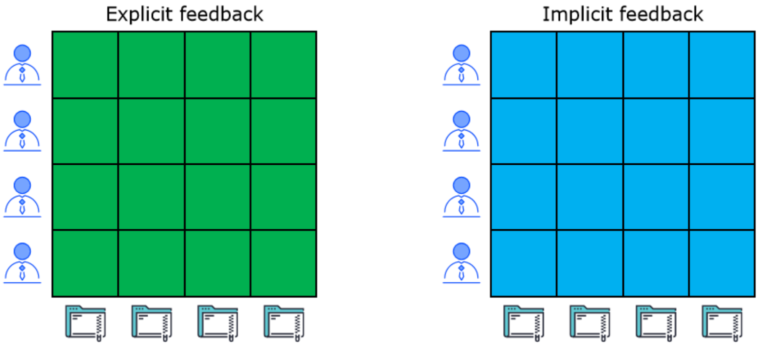

Figure 3: Explicit feedback and implicit feedback between users and projects.

© 2024 U-turn Press LLC, http://www.cad-journal.net

For each song, various features can be extracted from the audio signal. These characteristics can reflect the tone, rhythm, melody and other attributes of songs. For each user, learn their music preferences by analyzing their listening history and other data. Through user preference modeling, we can calculate each user's preference for different songs. According to the results of user preference calculation, the user's preference degree for different songs is taken as input, and the recommendation algorithm is used to output the songs that best meet the user's preference. User's music preference is influenced by short-term music interest and long-term music interest, so the interest preference of user's music can be expressed as:

$$D=\{M,N\} \tag{8}$$

In the formula: M represents short-term musical interest, and N represents long-term musical interest. Due to the variety of musical interests, M and N are respectively expressed as:

$$M=\{S_1,S_2,\ldots,S_n\} \tag{9}$$

$$N=\{L_1,L_2,\ldots,L_n\} \tag{10}$$

The user's music interest preference is expressed as:

$$U=\{S_1,S_2,\ldots,S_n,L_1,L_2,\ldots,L_n\} \tag{11}$$

Match the user's preference model with the music features, and determine the music works that users may be interested in by calculating the similarity or probability value. According to the matching results, the music works that users may be interested in are generated into a recommendation list according to certain sorting rules. According to the user's feedback on the music in the recommendation list, the recommendation algorithm is adjusted and optimized to continuously improve the accuracy of recommendation and user satisfaction.

5 EXPERIMENTAL RESULTS AND ANALYSIS

5.1 Recommended Performance Test

Amazon data set is a large data set containing a large quantity of data resources, which consists of several different types of data sets, including text, pictures, audio, video and GIF animation. This section evaluates the performance of different recommendation algorithms in music recommendation system on Amazon data set, and compares DBN algorithm with traditional CF algorithm, matrix decomposition algorithm and K- nearest neighbor algorithm. By training these algorithms on the data set and evaluating the test set, the accuracy, recall and recommendation error of various algorithms can be calculated. In the system load test, the method of gradually increasing user requests is adopted to simulate the actual operation scenario, and the maximum load changes of each algorithm in different operation cycles are recorded. Through the system load test experiment, we can better understand the performance and stability of different algorithms in practical application. The comparison of music recommendation accuracy and recall of the algorithms is shown in Figure 4 and Figure 5, respectively.

The algorithm in this article performs well in accuracy. With the adjustment of the algorithm parameters, the accuracy is obviously improved. This may be because the DBN model can effectively learn and extract the intrinsic characteristics of music data, thus more accurately identifying and recommending music that meets users' interests. The algorithm also has a good performance in recall. The recall reflects the ability of the algorithm to find music that really meets the interests of users. The DBN algorithm has a high recall at a low threshold, which means that it can cover more related music and enrich the recommendation results. With the increase of the threshold, the recall decreased, but remained within an acceptable range. On the whole, DBN algorithm shows high accuracy and good recall on Amazon data sets. This result shows that the proposed computer-aided MIR and recommendation algorithm is effective and practical in dealing with large-scale and high-dimensional music data.

© 2024 U-turn Press LLC, http://www.cad-journal.net

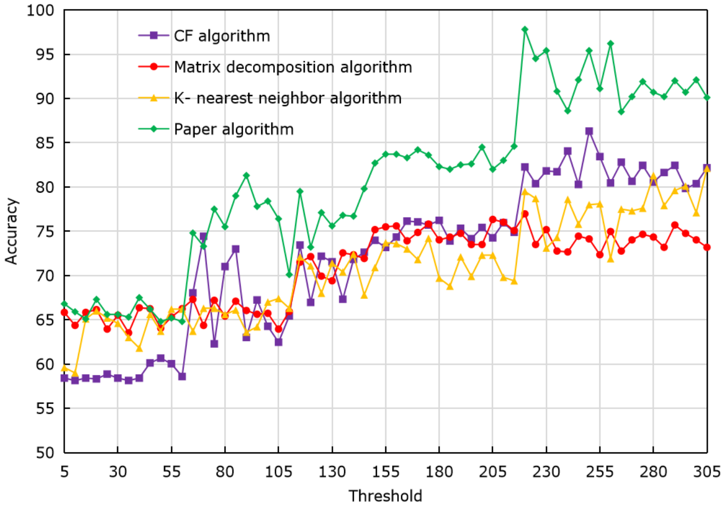

Figure 4: Recommended accuracy test.

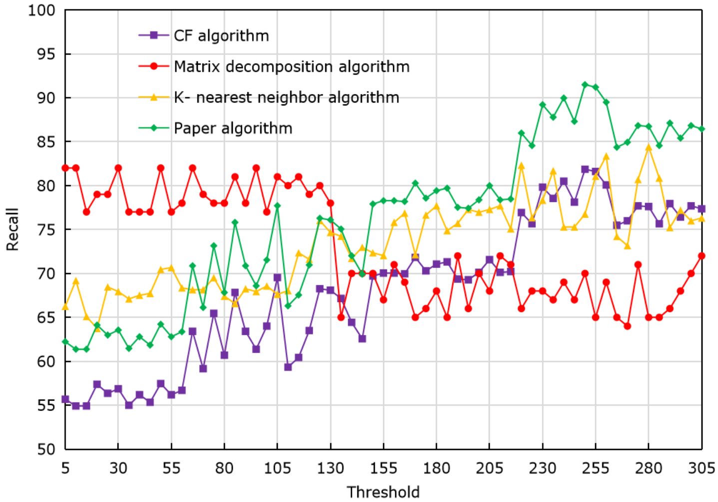

Figure 5: Recommended recall test.

From Figure 6, we can see that the DBN algorithm used in this article is compared with the traditional algorithms including CF algorithm, matrix decomposition algorithm and K- nearest neighbor algorithm in terms of music recommendation error.

On the whole, DBN algorithm shows better performance than traditional algorithms in music recommendation error. When the error threshold is low, the error rate of DBN algorithm is obviously lower than that of various traditional algorithms. This means that when recommending music, using DBN algorithm can predict users' music preferences more accurately. With the increase of error threshold, although the error rate of DBN algorithm has increased, it still remains at a low level and is better than the traditional algorithm. At low error threshold, the error rate of CF algorithm is low, but at high error threshold, the error rates of matrix decomposition algorithm and K- nearest neighbor algorithm are relatively low. This means that different traditional algorithms may have different advantages and application scope in different scenarios.

© 2024 U-turn Press LLC, http://www.cad-journal.net

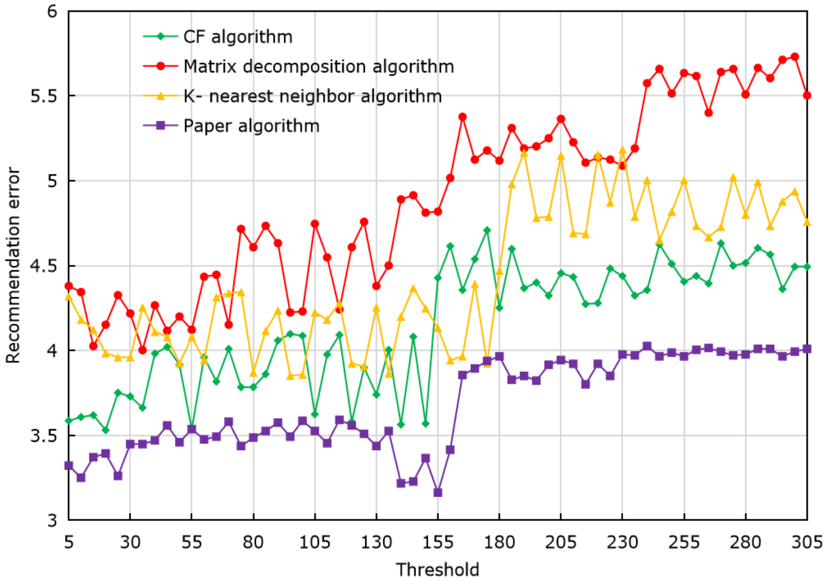

Figure 6: Algorithm error test.

Compared with traditional algorithms, DBN algorithm has better performance in dealing with music recommendation tasks. This is mainly due to the fact that DBN algorithm can automatically capture the complex features and high-level abstract concepts of music data, and has strong generalization ability.

5.2 System Load Test

For any recommended system, the load capacity of the system is one of the important indexes to measure its performance. The stronger the load capacity of the system, it shows that the system can better cope with and handle a large quantity of user requests and data, and provide users with more efficient services. IoT recommendation systems for different algorithms will face different challenges and limitations in the test. For example, some algorithms may pay more attention to accuracy and precision, so they may need more computing resources and time when processing a large quantity of data; Other algorithms may pay more attention to real-time and response speed, so they may need more concurrent processing ability and memory bandwidth when facing a large quantity of requests. Then, the system load test is carried out, and the maximum load change of IoT recommendation system combined with different algorithms in different operation cycles is tested. In the system load test, the stability and performance of the IoT recommendation system with different algorithms can be understood by observing the maximum load changes in different operating cycles. Figure 7 shows the results of the system load test.

CF algorithm and algorithm based on matrix decomposition show good performance and scalability when dealing with a large quantity of user requests and data. K- nearest neighbor algorithm also has high efficiency and scalability when dealing with large-scale data, but the post-processing time has increased. DBN algorithm shows a relatively high load level in the whole running cycle, and has the ability to automatically learn data features and abstract concepts, which is suitable for processing complex and large-scale music data. In the initial operation period, the load of the algorithm increased rapidly, but with the passage of time, the load growth rate gradually slowed down and stabilized. DBN algorithm has the ability to automatically learn data features and abstract concepts, which makes it more efficient and accurate when dealing with complex and large-scale music data.

© 2024 U-turn Press LLC, http://www.cad-journal.net

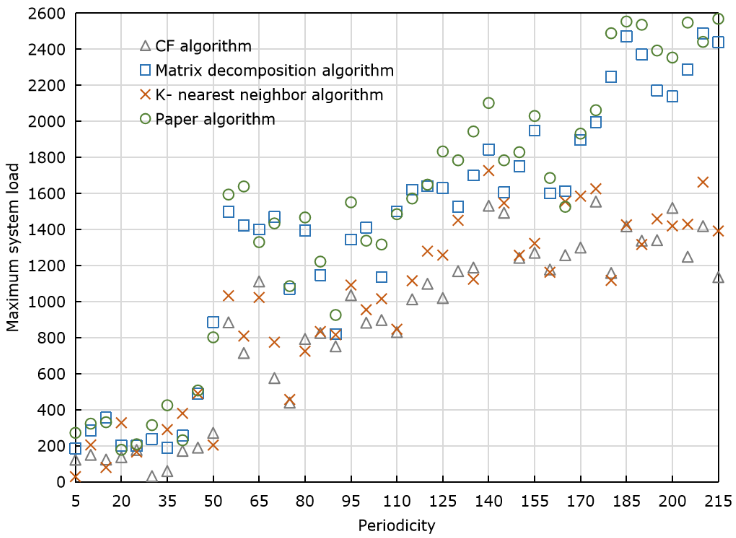

Figure 7: Load situation of the system.

6 CONCLUSIONS

With the growth of network technology, the number and types of music resources have multiplied, making it more difficult for people to find the music they need. IoT not only facilitates people's life, but also brings new features such as data diversity, dynamics, heterogeneity and real-time, which brings new challenges to information retrieval technology. This article proposes a DL-based MIR and IoT recommendation system, which aims to solve the challenges brought by the diversity, dynamics, heterogeneity and real-time performance of IoT data and provide more accurate and efficient computer-aided MIR and recommendation services. By comparing with the traditional recommendation algorithm, the advantages of this algorithm in recommendation accuracy, recall and real-time performance are verified. Through the system load test experiment, the maximum load changes of different algorithms in different operation cycles are evaluated. The implementation of this algorithm can provide useful reference and support for the design and optimization of IoT recommendation system, improve the performance and stability of recommendation system, and thus promote the sustainable growth of music industry. In recommendation system, user's feedback and behavior are very important data sources. How to analyze and use these data more deeply in order to improve the existing recommendation algorithm and improve the recommendation effect is an important research direction.

Hongxia Liu, https://orcid.org/0000-0003-1730-2972
Chenjing Zhao, https://orcid.org/0009-0002-0202-2564

REFERENCES

[1] Ahmad, M.-S.; Shah, S.-M.: Mitigating malicious insider attacks in the internet of things using supervised machine learning techniques, Scalable Computing: Practice and Experience, 22(1), 2021, 13-28. https://doi.org/10.12694/scpe.v22i1.1818

[2] Aldhyani, T.-H.; Alkahtani, H.: Artificial intelligence algorithm-based economic denial of sustainability attack detection systems: cloud computing environments, Sensors, 22(13), 2022, 4685. https://doi.org/10.3390/s22134685

© 2024 U-turn Press LLC, http://www.cad-journal.net

[3] Alkahtani, H.; Aldhyani, T.-H.: Developing cybersecurity systems based on machine learning and deep learning algorithms for protecting food security systems: industrial control systems, Electronics, 11(11), 2022, 1717. https://doi.org/10.3390/electronics11111717

[4] Alnaim, A.-K.; Alwakeel, A.-M.: Machine-learning-based IOT–edge computing healthcare solutions, Electronics, 12(4), 2023, 1027. https://doi.org/10.3390/electronics12041027

[5] Bout, E.; Loscri, V.; Gallais, A.: How Machine Learning changes the nature of cyberattacks on IoT networks: A survey, IEEE Communications Surveys & Tutorials, 24(1), 2021, 248-279. https://doi.org/10.1109/COMST.2021.3127267

[6] Carpentier, S.-M.; McCulloch, A.-R.; Brown, T.-M.; Faber, S.-E.; Ritter, P.; Wang, Z.; McIntosh, A.-R.: Complexity matching: brain signals mirror environment information patterns during music listening and reward, Journal of Cognitive Neuroscience, 32(4), 2020, 734-745. https://doi.org/10.1162/jocn_a_01508

[7] Dake, D.-K.: Artificial Intelligence Self-Organizing (AI-SON) Frameworks for 5G-Enabled Networks: A Review, Journal of Computer and Communications, 11(4), 2023, 33-62. https://doi.org/10.4236/jcc.2023.114003

[8] Dang, H.-V.; Tatipamula, M.; Nguyen, H.-X.: Cloud-based digital twinning for structural health monitoring using deep learning, IEEE Transactions on Industrial Informatics, 18(6), 2021, 3820-3830. https://doi.org/10.1109/TII.2021.3115119

[9] Gyamfi, E.; Jurcut, A.: Intrusion detection in internet of things systems: a review on design approaches leveraging multi-access edge computing, machine learning, and datasets, Sensors, 22(10), 2022, 3744. https://doi.org/10.3390/s22103744

[10] Hernandez, O.-C.; Zay, P.-I.; Hernandez, L.-C.; Beltran, J.-R.: A comparison of deep learning methods for timbre analysis in polyphonic automatic music transcription, Electronics, 10(7), 2021, 810. https://doi.org/10.3390/electronics10070810

[11] Huang, Z.; Jia, X.; Guo, Y.: State-of-the-art model for music object recognition with deep learning, Applied Sciences, 9(13), 2019, 2645. https://doi.org/10.3390/app9132645

[12] Imran, G.-Z.; Alshahrani, A.; Fayaz, M.; Alghamdi, A.-M.; Gwak, J.: A topical review on machine learning, software defined networking, internet of things applications: Research limitations and challenges, Electronics, 10(8), 2021, 880. https://doi.org/10.3390/electronics10080880

[13] Lozano, M.-Á.; Jiménez, B.-D.-M.; Valera, R.-A.; Santana, J.-F.; Moreno, G.-M.-N.: Context-aware recommender systems in the music domain: A systematic literature review, Electronics, 10(13), 2021, 1555. https://doi.org/10.3390/electronics10131555

[14] Mao, Z.; Liu, X.; Peng, M.; Chen, Z.; Wei, G.: Joint channel estimation and active-user detection for massive access in Internet of Things—A deep learning approach, IEEE Internet of Things Journal, 9(4), 2021, 2870-2881. https://doi.org/10.1109/JIOT.2021.3097133

[15] Pandit, M.-K.; Mir, R.-N.; Chishti, M.-A.: Adaptive deep neural networks for the internet of things, International Journal of Sensors Wireless Communications and Control, 10(4), 2020, 570-581. https://doi.org/10.2174/2210327910666191223124630

[16] Pei, Z.; Wang, Y.: Analysis of computer aided teaching management system for music appreciation course based on network resources, Computer-Aided Design and Applications, 19(1), 2021, 1-11. https://doi.org/10.14733/cadaps.2022.s1.1-1

[17] Rajyalakshmi, V.; Lakshmanna, K.: A review on smart city-IoT and deep learning algorithms, challenges, International journal of engineering systems modelling and simulation, 13(1), 2022, 3-26. https://doi.org/10.1504/IJESMS.2022.122733

[18] Ramírez, J.; Flores, M.-J.: Machine learning for music genre: multifaceted review and experimentation with audioset, Journal of Intelligent Information Systems, 55(3), 2020, 469-499. https://doi.org/10.1007/s10844-019-00582-9

[19] Sun, Y.; Nurellari, E.; Ding, W.; Shu, L.; Huo, Z.: A partition-based mobile-crowdsensing-enabled task allocation for solar insecticidal lamp internet of things maintenance, IEEE Internet of Things Journal, 9(20), 2022, 20547-20560. https://doi.org/10.1109/JIOT.2022.3175732

© 2024 U-turn Press LLC, http://www.cad-journal.net

[20] Zamee, M.-A.; Han, D.; Won, D.: Online hour-ahead load forecasting using appropriate time-delay neural network based on multiple correlation–multicollinearity analysis in IOT energy network, IEEE Internet of Things Journal, 9(14), 2021, 12041-12055. https://doi.org/10.1109/JIOT.2021.3133002

© 2024 U-turn Press LLC, http://www.cad-journal.net