

## Research Article

# Multisource Heterogeneous Data Fusion Analysis of Regional Digital Construction Based on Machine Learning

Mengmeng Jiang,<sup>1</sup> Qiong Wu ,<sup>1</sup> and Xuetao Li<sup>2</sup>

<sup>1</sup>School of Economics, Wuhan Donghu University, Wuhan, 430212 Hubei, China

<sup>2</sup>School of Economics and Management, Hubei University of Automotive Technology, Shiyan, 442000 Hubei, China

Correspondence should be addressed to Qiong Wu; wuqiong@wdu.edu.cn

Received 13 August 2021; Revised 14 October 2021; Accepted 2 November 2021; Published 10 January 2022

Academic Editor: Mu Zhou

Copyright © 2022 Mengmeng Jiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In modern urban construction, digitalization has become a trend, but the single source of information of traditional algorithms can not meet people's needs, so the data fusion technology needs to draw estimation and judgment from multisource data to increase the confidence of data, improve reliability, and reduce uncertainty. In order to understand the influencing factors of regional digitalization, this paper conducts multisource heterogeneous data fusion analysis based on regional digitalization of machine learning, using decision tree and artificial neural network algorithm, compares the management efficiency and satisfaction of school population under different algorithms, and understands the data fusion and construction under different algorithms. According to the results, decision-making tree and artificial neural network algorithms were more efficient than traditional methods in building regional digitization, and their magnitude was about 60% higher. More importantly, the machine learning-based methods in multisource heterogeneous data fusion have been better than traditional calculation methods both in computational efficiency and misleading rate with respect to false alarms and missed alarms. This shows that machine learning methods can play an important role in the analysis of multisource heterogeneous data fusion in regional digital construction.

## 1. Introduction

With the rapid development of China's urbanization process, people's demand for a better life is increasing, and the traditional urban management model can no longer meet the needs of reality. At the same time, the rapid development of modern information technology provides a new way for urban management. It applies modern information technology to city management, divides the city management area into a series of unit grids, and classifies its components and events. The platform quickly finds problems and solves them in a timely manner, improving the efficiency of urban management.

Many source data fusion is a method which deals with data from multiple sources. The estimates and judgments from the original data sources are derived in terms of conclusions attained and knowledge recognition developed to increase data

confidence, improve reliability, and reduce uncertainty. It is an idea that comes from the multisensor datalink technique proposed during the 1970s to obtain more complete and successful information and to improve the robustness and accuracy of the system because the detection range and computing power as well as the accuracy of each sensor has been limited. Similarly, in combat scenarios, there are potential conflicts due to the heterogeneity of single sensor information sources, the differentiation of multiple sources of information, and the high intensity of information sources and electronic countermeasures. As a result, multisensor information fusion technology was born and began to flourish [1].

Experts at home and abroad have done a lot of research on the digital construction of districts. Ding et al. use three-dimensional (3D) laser scanning to realize the RE process. The framework also incorporates supporting technologies (virtual reality, 3D printing, and prefabrication) to better

understand design and construction and tools (work breakdown structure and model breakdown structure) to improve the quality of organization and management. The implementation of this proposed framework in a refurbished shopping center has optimized the efficiency of the renovation process by 15%, reduced design changes by 30%, and rework by 25% [2]. Rytova and Gutman propose methods to address both tasks using modern tools, including balanced exponential systems adapted to regional development. He believed that the quality of innovative activities and human capital should be increased, creating effective conditional arrangements for regional development of socioeconomic systems, and innovative creation and development of the digital economy, and the formalization of strategic objective assessment using a fuzzy set method [3]. López et al. believe that digitalization is a global phenomenon that affects all human activities. The public administration department also incorporates new information and communication technologies into its structure and public sector. Computer auditing is a tool that allows auditing of public administration departments and improving accountability. The article examines the main advantages and risks of digitization and provides a case study of a regional audit institution that implements computer audits in Spain [4]. For data fusion solutions, there are also different views on this research. Queiroz et al. established the interregional DIH created by the cooperation between entities to transform the region into a reference pole for innovation. This article is aimed at describing the innovation quality management and improvement strategies developed by this cross-regional DIH within the scope of the DISRUPTIVE project. In addition to considering the personal strategies of affiliated members, it also includes the aspect of cooperation, on the sharing of knowledge, technology, and skills, aimed at improving the quality of innovation, and the adoption of digitalization by companies in the region [5]. Bareinboim and Pearl focus on the latest technology of sensor fusion technology, which is applied to sensors embedded in mobile devices as a means to help identify the daily activities of mobile device users. Sensor data fusion technology is used to integrate data collected from multiple sensors to improve the reliability of algorithms that identify different activities [6]. Liu and Huang believe that the rapid development of sensing and computing technology has enabled multiple sensors to be embedded in a system to simultaneously monitor the degradation state of an operating unit. It is aimed at solving these two challenging problems in a unified way. Specifically, a method was developed to construct a health index by fusing multiple degradation-based sensor data [7]. Mou et al. discussed the scientific results of the 2016 Data Fusion Competition organized by the Image Analysis and Data Fusion Technical Committee of the IEEE Society of Earth Sciences and Remote Sensing. These include a pair of ultrahigh resolution panchromatic and multispectral Deimos-2 images and videos taken by the Iris camera on the International Space Station. The problems solved and the technologies proposed by the contest participants covered a wide range of topics, as well as mixed ideas and methods from remote sensing, video processing, and computer vision [8]. These methods provide a certain

reference for this article, but because the amount of data in related studies is too small and the sample experiment time is short, it is difficult to reproduce the experimental conclusions.

A novel feature of this paper is the use of decision tree-based algorithms and artificial neural network algorithms which are used as generative methods to construct an algorithm for complex model discriminations. With respect to the uncertain information available in the data source, reliability of assignment of the method should be used to determine the confidence function. Finally, the production weight is determined through a reliable mathematical structure, and a weighted data fusion correction model is established to be better applied in practice. The experiment has achieved good results.

## 2. Regional Digital Construction Methods

*2.1. Data Fusion.* Change the data with new data from multiple data sources to draw the top choices. This book is superior to options made by a data source or a single decision book [9, 10]. The data model should change the breakpoint by changing the breakpoint. The data comes from a variety of information sources, which improves the final segmentation performance, thereby improving the quality of the final decision.

It reflects the many features of the information and the target, and only when the features are fully analyzed are they effectively consolidated [11]. Through the analysis of multi-source heterogeneous data fusion, the characteristics of multisource heterogeneous data fusion are summarized as follows:

*Heterogeneity:* the data for fusion processing in a data fusion system often comes from data generated by multiple independent systems. Because different data sources were independent of each other, not only the data model is heterogeneous but also the semantics and syntax of the data. There are also different degrees of heterogeneity.

*Scatterability:* in multisource heterogeneous data, which is often distributed heterogeneously, interference between some data is transmitted through the network, so there are problems of network transmission reliability and security, etc. As well as the timeliness and error of data, which can be caused by various interference problems, there are some obstacles to the data conversion system.

*Autonomy:* since the source of data may come from some independent systems, these mutually independent systems have strong autonomy, and they are likely to change their own structure and data according to their own needs without notifying the fusion system. This poses a great challenge to the data fusion system, and the robustness of the fusion system will also be affected.

It is observed that the main expressions and manifestations of heterogeneous data in converged systems can be divided into syntactic dissimilarities and semantic dissimilarities [12]. A grammatical heterogeneity mainly represents the fact that the same objects and facts in the domain are described in different ways. The different data table structures are for example different in different databases with

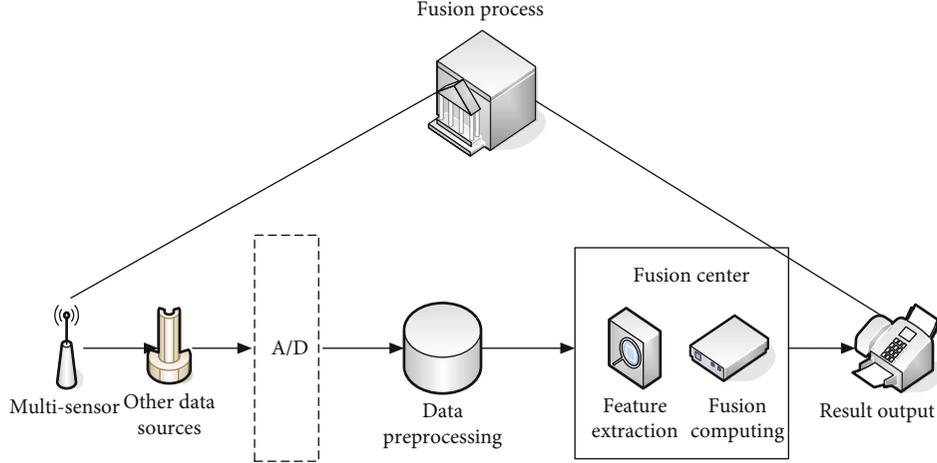


FIGURE 1: Multisource data fusion process.

the same data stored, due to the different naming rules and data types that are present in different databases [13]. For example, when describing parts of the system, some files use parts to describe, and some files use components to describe. Compared with the elimination of semantic heterogeneity, the elimination of grammatical heterogeneity only needs to implement the mapping relationship from field to field or record to record in specific operations and then resolve the naming and data type conflicts between these concepts. The heterogeneous data in the fusion system is shown in Figure 1:

For that reason, the syntactical heterogeneity which is easier to achieve, there is no reason to understand about the specific content and meaning of the data in the mapping process. That can be solved if we can obtain data structure in the information and complete the mapping in different data sources between the structure of the data [14]. Semantic heterogeneity is relatively more complicated than synthetic heterogeneity, because semantic heterogeneity usually destroys the personality between fields and needs to be processed immediately at the data content level.

**2.2. Data Processing.** An important guidance for data preprocessing operation can be found in the stream shape learning algorithm. From these analysis, this paper proposes a data preprocessing operation combining noise-removal operation and noise-cancellation operation in a streaming learning method, i.e., data preprocessing operation of the streaming learning algorithm [15, 16]. Located on a local linear was plane of the manifold structure. Each data point in the manifold structure has  $k$  neighbors.

$$Z_j = B^T(x_{ij} - d) \in R^m. \quad (1)$$

This transformation should be reversible, where  $Z_j$  is the mapping result of  $X_{ij}$ , which can also be obtained by  $Z_j$  doing the inverse transformation  $X_{ij}$ .  $d$  is the translation vector, and  $m$  is the spatial dimension after projection and satisfies the orthogonal normalization condition, namely

$$b_j^T b_k = \delta_{jk}, \quad (2)$$

where at that time,  $j = k$  and  $\delta_{jk} = 1$ ; otherwise,  $\delta_{jk} = 0$ .

$$X'_{ij} = d + BB^T(X_{ij} - d) = d + BZ_j. \quad (3)$$

The formula is inverse transformation. In actual operation, due to the influence of noise data or different transformation methods, there are errors between  $X'_{ij}$  and  $X_{ij}$  them, as shown below:

$$\varepsilon_j = X_{ij} - X'_{ij} = X_{ij} - d - BB^T(X_{ij} - d). \quad (4)$$

Perform objective optimization operations on  $B$  and  $d$ . The specific formula is as follows:

$$\min \left\{ \sum_{j=1}^k \omega_j \|\varepsilon_j\|^2 \right\} = \min \left\{ \sum_{j=1}^k \omega_j \|X - d1^T - BZ_j\|^2 \right\}, \quad (5)$$

where  $\omega_j$  is the weight of the error  $\varepsilon_j$ . According to the feature decomposition, the minimum weighted mean square value of  $B$  is obtained.  $S$  is the weighted covariance matrix of neighboring points. Then, calculate the minimum mean square value of the translation vector  $d$ :

$$d = \frac{\sum_{j=1}^k \omega_j X_{ij}}{\sum_{j=1}^k \omega_j}, \quad (6)$$

$$S = \frac{1}{k} \sum_{j=1}^k \omega_j (X_{ij} - d)(X_{ij} - d)^T. \quad (7)$$

When the above formula is transformed, the size is related. The weight of the sample point reflects the possibility that the point  $\varepsilon$  is noisy data. If the error is large, it means that the point is likely to be noised; otherwise, the point is less likely to be noise. The following functional relationship is satisfied between the weight and the error:

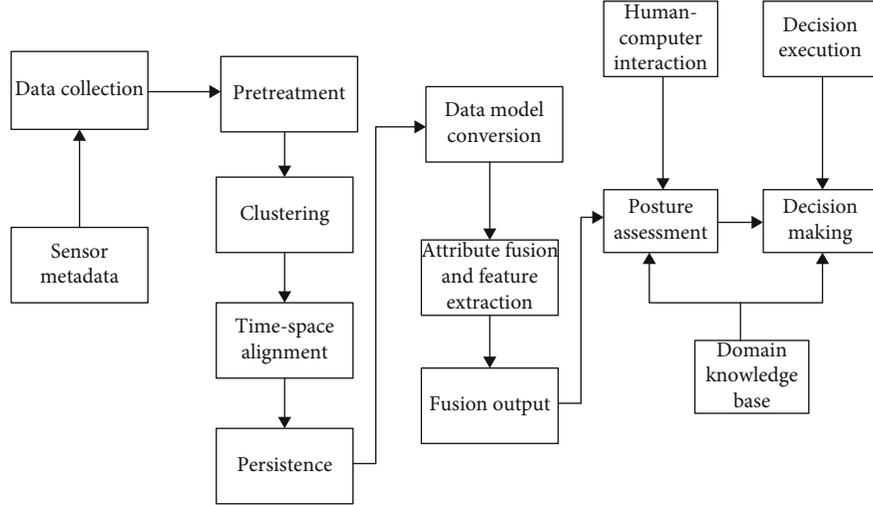


FIGURE 2: Multisensor data fusion model.

$$\omega(\varepsilon) = \begin{cases} 1, & |\varepsilon| \leq \mu, \\ \exp\left[-\frac{(\varepsilon - \mu)^2}{2\sigma^2}\right], & \text{otherwise} \end{cases} \quad (8)$$

$\varepsilon$  and  $\mu$  defined as follows:

$$\mu = \frac{1}{n} \sum_{i=1}^n \varepsilon_i, \quad (9)$$

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (\varepsilon_i - \mu)^2. \quad (10)$$

Divide the data  $X_i$  to be classified into  $c$  fuzzy partition groups, find the cluster center under each fuzzy partition, initializing a membership  $i * j$  matrix  $U$  with a random number in the range of  $[0, 1]$ , and any element  $u_{ij}$  in the matrix satisfies the condition:

$$\sum_{i=1}^c u_{ij} = 1, \forall i = 1, 2, \dots, n. \quad (11)$$

$u_{ij}$  represents the degree  $X_i$  of membership of the cluster centers  $v_i$  and calculates each cluster center  $v_j$ :

$$v_j = \frac{\sum_{i=1}^n 1u_{ij}^m x_i}{\sum_{i=1}^n 1u_{ij}^m}. \quad (12)$$

Calculate the cost function. If the cost function is less than a certain threshold or the change of the cost function during two iterations is less than a certain threshold, the algorithm stops  $\beta$ , and the cost function is

$$J(U, v_1, \dots, v) = \sum_{i=1}^c J_i = \sum_{i=1}^n u_{ij}^m d^2(x_i, v_j). \quad (13)$$

Update the membership matrix  $U$ , and then return to the step:

$$u_{it} = \sum_{j=1}^c (d_{it}/d_{jt})^{-2/(m-1)}. \quad (14)$$

For the membership matrix output by the algorithm  $U$ , no human intervention is required in the algorithm implementation process. In order to avoid the possible misjudgment of this method, based on the cosine similarity, the cosine value of the angle between the point and the cluster center is used to weight the Euclidean distance. But

$$d_v(x_t, v_j) = \text{sim}(x_t^{(j)}, v_j) * \sqrt{(x_t^{(j)} - v_j)^T (x_t^{(j)} + v_j)}. \quad (15)$$

Among them  $t = |v_j|$ ,  $|v_j|$  represents the number of samples in a cluster  $v_j$  that is the cluster center, and  $x_t^{(j)}$  represents all sample points  $v_i$  in the cluster where the cluster center is located.

During the feature-level data fusion stage, which is the input from the structured data file output at the data-level data fusion stage, it is necessary to describe the sensor data in a unified manner, i.e., transform the data model, with a view to facilitating feature extraction and association of browser data, while at the same time sharing the sensor data with the subsequent fusion process [17, 18]. The multisensor data fusion model is shown in Figure 2:

The decision-level data fusion stage accepts situation reports from the feature-level data fusion stage. These situation reports are only a situation description of a certain aspect of the current scene. In order to obtain the overall situation in the current scene, these situation ontology fragments need to be fused and fused. After that, the situation ontology has a global situation description for the current scene, and then, the situation assessment can be carried out with the help of the global situation ontology

description. Situation assessment is a higher-level abstraction of data [19]. An eventual decision would be generated based on the current situation assessment then the decision information in the global situation ontology about the current situation by searching in the global situation ontology on the basis of the rule-based ontology reasoning [20].

**2.3. Data Model Conversion in Data Fusion.** The main task of the multisensor feature-level data fusion stage is to deal with the problem of information representation and correlation [21]. Feature-level data fusion needs to extract the attribute information of the fusion target, but the data description form of different sensors is different, so it is difficult to extract attributes from the data of different description forms, which requires a unified and standardized description of these data. This is the operation of data model conversion performed in the feature-level data fusion stage [22]. The data model conversion operation is aimed at shielding the obstacles to data fusion caused by the difference in data description form, so that the sensor data can be shared and fused in different fusion stages in the same representation form [23].

In the environment of multisensor data fusion, sensor data needs to be processed at different fusion levels, and different data representations and different data abstraction levels need to design corresponding data processing methods [24]. In order to better achieve multisource heterogeneous data fusion, we have drawn the data model conversion flowchart of multisource heterogeneous data fusion, as shown in Figure 3:

As this flow diagram shows, the sensor metadata at the sensor's metadata layer is captured for the data level fusion layer for organized representation, where various structured data files are formed at the data level and stored in different media [25]. When some sensor data needs to be processed at a higher level, heterogeneous data in different formats need to be processed at the data level to form structured data about the fusion target and upload it to a higher level of data processing. After the feature-level data fusion layer receives the structured data file about the fusion target input by the data-level fusion layer, in order to realize the feature extraction of the fusion target and the unified processing of heterogeneous data, these heterogeneous data need to be described uniformly [26].

Since the sensor is susceptible to various interferences from the external environment, the data obtained from the sensor is likely to be distorted, and the error registration operation is performed on these data in the preprocessing stage [27]. The data after the preprocessing stage is of higher quality than the original data, and it is also convenient for data clustering operations. Data clustering is to classify and synthesize multiple vectors describing the same attribute of the target, so that the sensor has a unified attribute description for a single attribute of the target, and then perform spatio-temporal alignment operations on these attribute descriptions. Because some data forms are not structured, it is difficult to establish an effective description of these data. In order to make inference judgments and combined decision-making for multisource heterogeneous data during

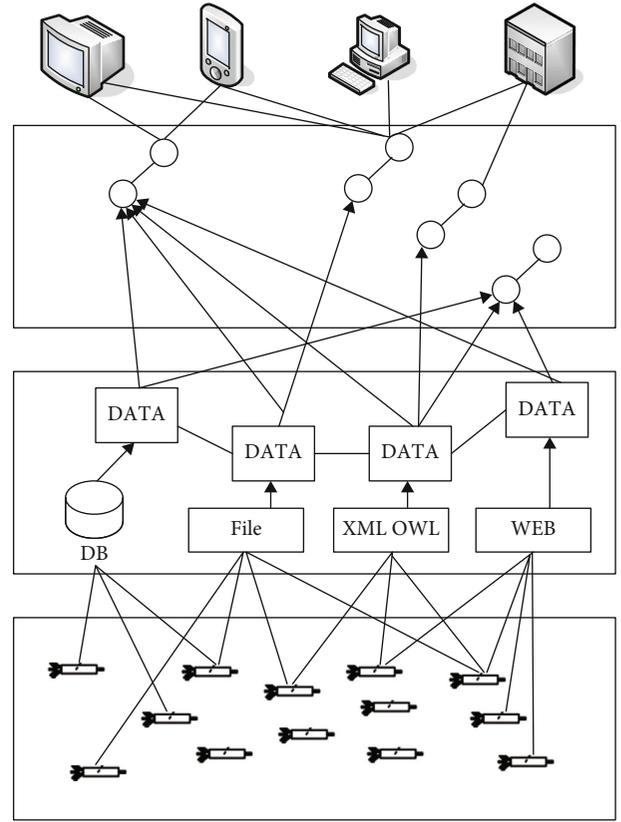


FIGURE 3: Data model conversion flowchart.

TABLE 1: Experimental equipment parameter list.

Category	Parameter
Server	Intel(R)
CPU	Core i7-7700
Main frequency	4.0 GHz
RAM	8 G
Operating system	Linux
Simulation software	MATLAB 2016
Cloud environment	Hadoop

TABLE 2: Basic information of the data set.

Data set	Training data set	Test data set
Total number of samples	494317	335287
Number of normal samples	95632	57482
Number of abnormal samples	245632	258037
Number of DoS samples	380334	218742
R2L sample number	1015	15078
Number of U2R samples	41	217
Probe sample number	4096	4055
Unknown number of samples	—	17601

fusion, the solution is to extract the semantics of multisource heterogeneous data in the feature-level data fusion stage and establish a unified data description [28].

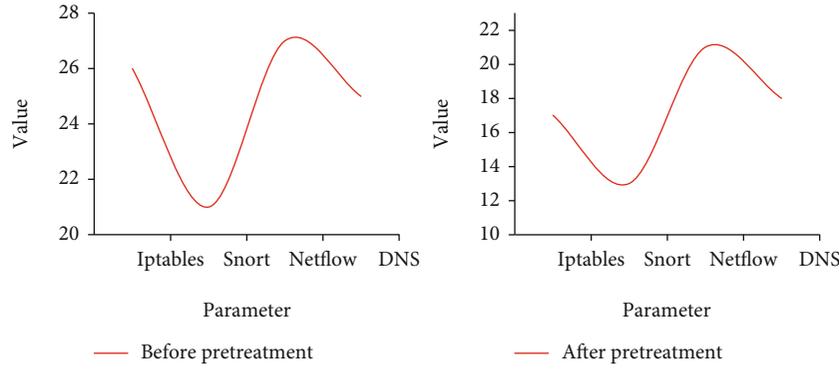


FIGURE 4: Data preprocessing results.

### 3. Regional Digital Construction Experiments and Conclusions

**3.1. Regional Digital Environment.** This article first builds a cloud environment to verify the machine learning algorithm proposed in the paper. We select the decision tree in machine learning and artificial neural network for comparison. According to the existing equipment in the laboratory and the preliminary scientific research work, the machine learning safe data source is selected for the experiment. The experimental equipment parameters are shown in Table 1.

The training data is shown in Table 2.

The data volume is reduced after the data source is pre-processed. In this experiment, a total of 4 GB of data was used for preprocessing, and the results of data preprocessing are shown in Figure 4.

The differences before and after data were large, as can be seen from Figure 4, where the maximum value before data processing reached 27, but the maximum value after processing was 22, with Iptables decreasing by about 20% after processing, and in the case of DNS decreasing by about 30%.

In order to comprehensively evaluate the decision tree and artificial neural network algorithm proposed by the theory, the proposed method is verified and analyzed through the existing classic data sources, and the comparison is made. We use the evaluation index of algorithm stability; it is associated with accuracy rate, mistake rate, and the missing rate. Implementation of the algorithm enhancement effect involves two stages: data preprocessing stages and security analysis. Verification metrics are determined by the classification results of the algorithm. When the input sample type is malicious and the detected result type is also malicious, it is called true positive (TP). When the sample is malicious and the result is normal, it is called false negative (FN). When the sample is normal and the result is normal, it is called true negative. Among them, “1” means normally, and “0” means error. The test results are shown in Table 3.

To compare the effect of the decision tree and the artificial neural network, we compare the differences between the training set and the test set between the error report rate and the missed report rate. Figure 5 shows the training set comparison result, and Figure 6 is the test set comparison result.

TABLE 3: Test results.

Index	TN	FN	FP	TP
Sample type	1	0	1	0
Test result	1	1	0	1

It is evident from the comparison of the two datasets from this paper that the overall performance of the proposed machine and the improved method KNN-DST based for D-S evidence theoretical is significantly better than that of the other method KNN-DST. More specifically, in the training set, because the data are all known data of this type, the difference between the two algorithms in terms of detection accuracy, false alarm rate, and disappearance rate is not obvious.

**3.2. Regional Digitization.** During the time-consuming comparison of different models, this paper compares the prediction time required to handle the number of events at different sizes, with the time-consuming comparison results in seconds as shown in Figure 7.

It can be seen that the prediction model processing with different orders of magnitude, the prediction time consumption is always lower than the traditional prediction model. With the amount of data as 30,000 points, the time consumption began to grow rapidly. The traditional model has the fastest increase in time-consuming when the processing volume is less than 5,000, and the growth tends to be stable when the data volume is 5,000 to 30,000, and it grows rapidly again when the processing data level is more than 30,000. Compared with traditional models, the heterogeneous data matching algorithm based on decision tree and artificial neural network in this paper is more accurate and less time-consuming.

For regional digital construction, we take a school in this city as an example to compare the digital proportions in this area for statistics. The specific results are shown in Figure 8.

It can be seen from Figure 8 that the digital construction of this area, taking our school as an example, has not changed much in recent years. In terms of teaching mode and network capacity, the digital construction of this school is basically stable at around 2, which basically does not meet the standards of digital construction. High-speed network

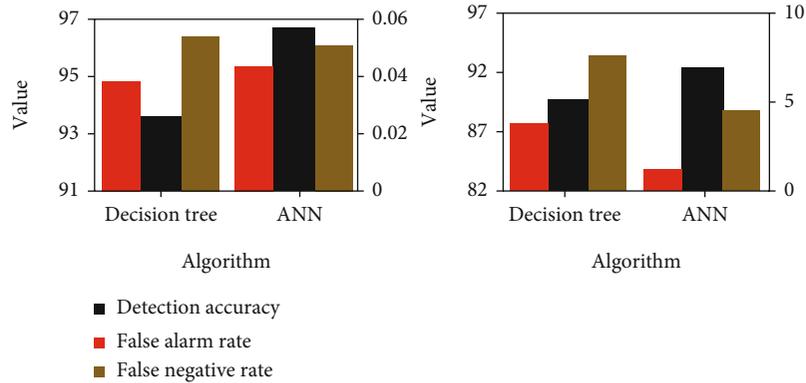


FIGURE 5: Training set comparison.

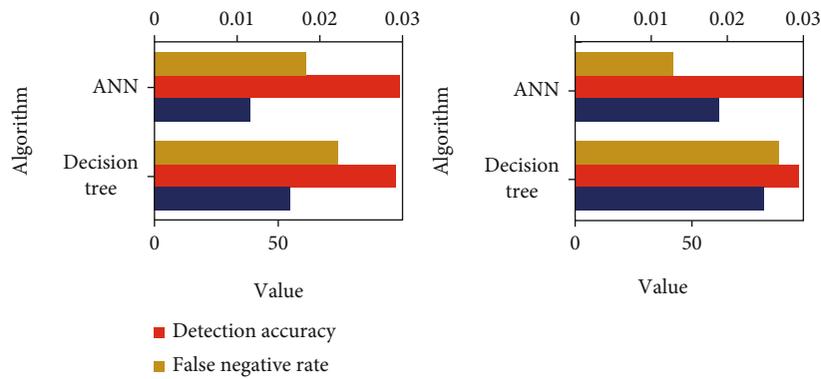


FIGURE 6: Test set comparison.

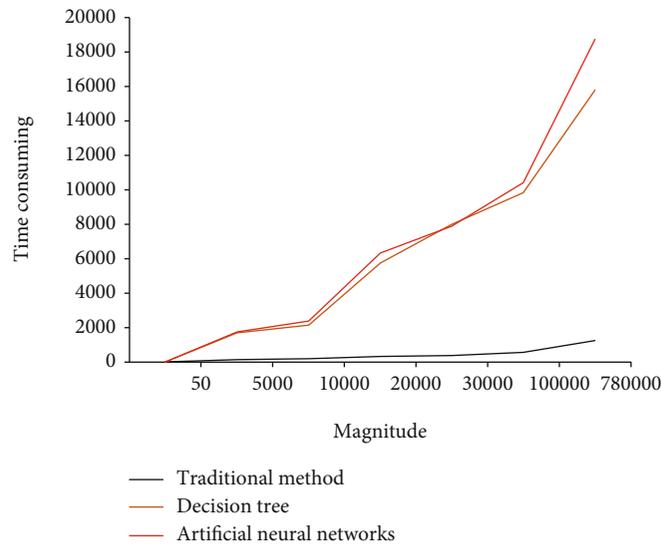


FIGURE 7: Model time-consuming comparison.

technology spans the distance of time and space. Teachers can teach a course at their own convenient time and place and guide students' learning without being restricted by the time and place of class. Therefore, we carried out a data fusion transformation for the digitization of the school, and the digital construction after the transformation is shown in Figure 9.

It can be seen that after the analysis of heterogeneous data fusion, the school's digital construction has shown an increasing trend. The average value of digital construction is about 3.2, which is an increase of about 1.2 compared to before the fusion, and the growth rate is about 60%. This shows that there has been a considerable increase in regional digital construction of data fusion through machine

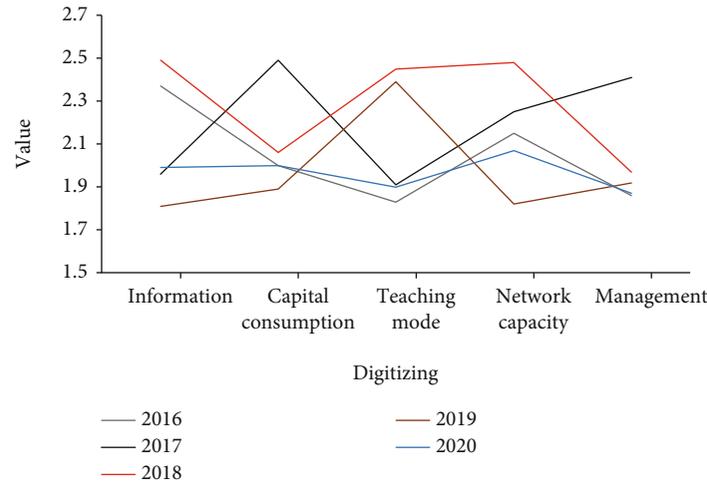


FIGURE 8: Digitization ratio statistics.

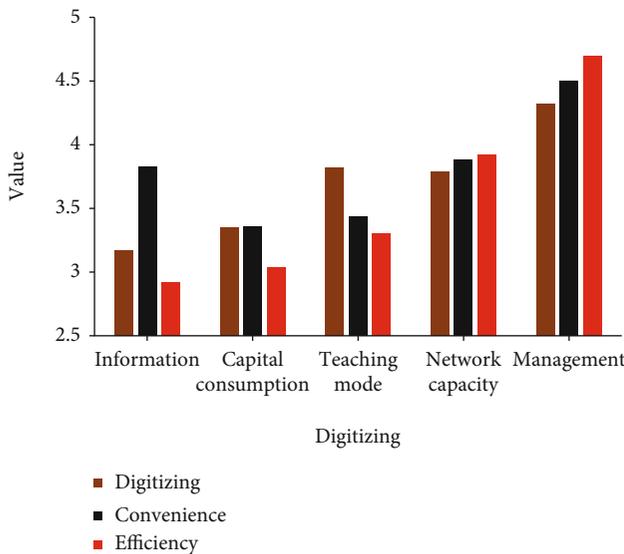


FIGURE 9: Digital construction parameters after integration.

learning. In order to compare the impact of digital construction in different regions on the region, we compare the management of the school before and after the integration, as shown in Figure 10.

It can be seen from Figure 10 that in the school management before data fusion, the school’s management level of teachers and students is at a low level. After problems occur in learning, the feedback time is longer and cannot be handled well. But after the data fusion processing, the management of the school has been greatly improved, the average value has been increased from 2.2 to 4.5, and the feedback of the school’s problems can also be dealt with in a timely manner. In order to verify the effectiveness of the data fusion, we select 10 days each to calculate the processing efficiency of different methods, and the results are shown in Figure 11.

It can be seen from Figure 11 that the efficiency of school transaction processing before and after the fusion is obvious. In the selected time, the comparison between the traditional

method and the decision tree and neural network selected in this paper is obvious. The two algorithms of the machine school are more effective in transaction processing. The efficiency is significantly better than traditional methods. Of course, in management, we should not only look at the resolution of affairs but also take into account personal feelings. Therefore, we conducted random surveys at the school to compare satisfaction under different methods. The results are shown in Figure 12.

It can be seen from the figure that there is a huge difference in satisfaction under different methods. The traditional management method is rude, without skill, and cannot understand the needs of students and teachers. As a result, it is still not satisfactory after the problem is solved. Under the machine school, after data fusion analysis, problems can be solved well, and people’s satisfaction can be improved.

#### 4. Discuss

*4.1. Impact of Digital Construction.* With the widespread application of the Internet of Things, the fusion technology of multisource heterogeneous data has gradually become a research hotspot in the field of data processing. Due to the grammatical and semantic heterogeneity of multisource heterogeneous data, the interaction, sharing, and fusion reasoning between data are facing obstacles, and the value contained in the data is difficult to be fully utilized. By using the ontology as the description model of multisensor data, the sensor data can be described uniformly, which is convenient for the fusion system to retain the original semantics of the sensor data during the data fusion operation at different fusion levels. The interfusion reasoning provides high-quality data.

Regional digital construction can effectively improve the management efficiency in the region. This is mainly due to the new management model. According to the actual situation of urban management, a management model different from other urban areas is adopted, and the urban management coordination mechanism is rebuilt to improve the

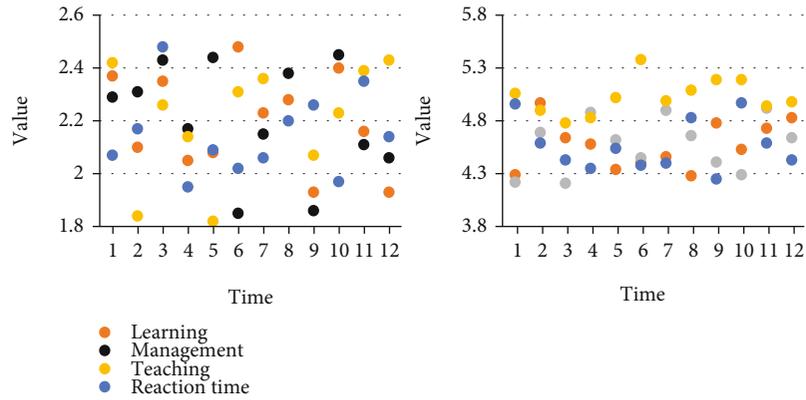


FIGURE 10: Postintegration learning management.

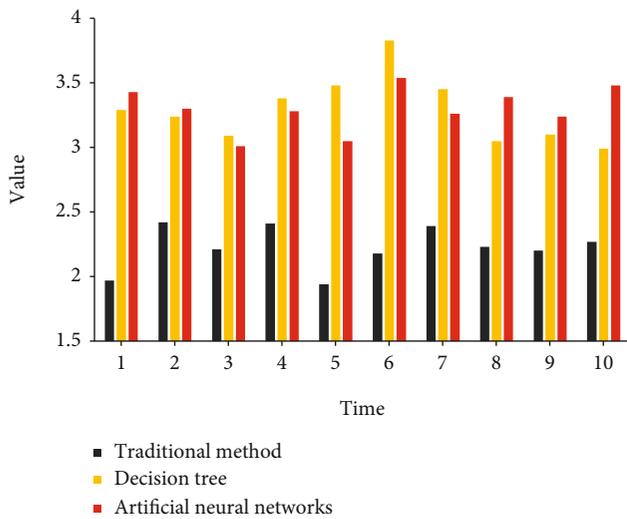


FIGURE 11: Efficiency comparison before and after fusion.

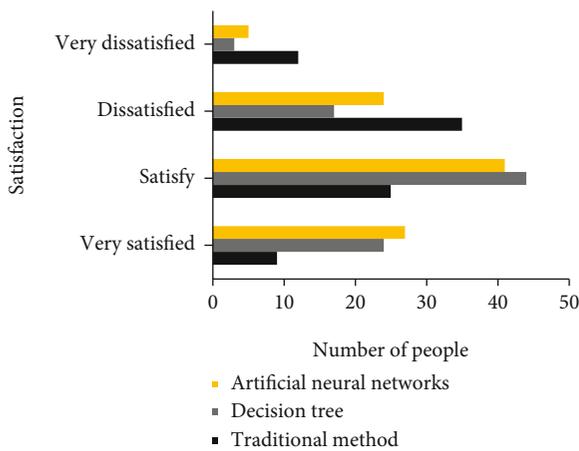


FIGURE 12: Satisfaction comparison survey.

urban management digital platform supervision center. With the construction of the command center, it can clarify the division of responsibilities and highlight the city management method of departmental cooperation and coordination. And it can also determine the work system of territorial

management, effectively break through the traditional city management model, develop from the pyramid level structure to flattening. At the meanwhile, it can improve city management work supervision and management, communication, and coordination capabilities, to achieve an optimized model to promote efficiency improvement.

In the digital construction of city management, make full use of the existing platform construction and information resources, do a good job of horizontal resource integration while doing vertical integration, and connect the platforms of other city-level city management-related departments and the platforms of subordinate district-level departments. It can effectively avoid duplication of construction to achieve resource sharing. In the newly constructed digital platform, high-level configuration and reserved interfaces can be avoided, which can avoid the inability to expand in later construction. It can be found that by integrating the resources of the original urban management-related departments and configuring new data resources at a high level, the urban management digital system between the city-level departments and the urban two-level departments can be interconnected, resource-sharing, and step-by-step. Consistently, how to improve the processing capacity and resource was allocation of urban management digitization in the quick and efficient way.

Digitization accelerates the process of information construction in universities by promoting and deepening information production. It will impact old teaching concepts and bring about reforms in teaching methods. In traditional educational thought, teachers are the main tools of teaching. In most cases, they passively accept various knowledge impart by teachers. However, with the development of information technology, the subject of educational activities will be changed. Teachers are more drivers, encouraging students to actively acquire relevant knowledge through digital platforms and resource networks and pay more attention to self-study. It supports students' learning initiative.

**4.2. Enlightenment of Digital Construction.** In the regional digital construction, through digital means, the problems in urban management are quickly collected, and the collected data is professionally analyzed. Through the analysis of the causes of urban management problems by

professionals, the city can be continuously optimized. The process of management improves the quality of service and avoids the recurrence of the same problem or the inadequate handling of one problem multiple times. At the same time, it opens the handling process, increases the enthusiasm of the public to participate, and creates an atmosphere for the government, society, and the public to work together in urban management.

During the promotion of digital construction, it is necessary to pay attention to the publicity and popularization of laws and regulations, so that the public can develop a clearer understanding of the relevant requirements of urban management, while strict law enforcement is carried out under adequate publicity and training, to establish the majesty of the law, to well apply legal weapons for urban management, while increasing the public's legal awareness on the establishment of a society in accordance with the rule of law. Each department has a clear division of functions and clear responsibilities. At the same time, each department holds regular meetings to communicate while using digital means to strengthen contact with the public. Finding shortcomings and making timely improvements can greatly improve the efficiency of urban management and avoid departmental interference. Refuse each other, misplaced cognition between the department and the public, and the information communication is not smooth. Through efficient communication, more resources can be provided to better provide city management services.

## 5. Conclusions

With the advent of the information explosion, data from a single source can no longer meet people's needs for data richness, real-time, accuracy, and reliability. This requires data fusion technology to obtain estimates and decisions from multisource data to improve the credibility of data; this paper uses heterogeneous data aggregation technology to study the regional digital design based on machine learning methods, and from an experimental point of view, good results have been achieved. At present, the development of evidence theory is mainly related to the structure of machine learning, and computer complexity is rarely considered. The next step should be to combine the construction of fast algorithms, construct similar fast algorithms for specific application areas, and also combine corresponding improved synthesis rules to improve the fusion efficiency of algorithms in practical applications.

## Data Availability

No data were used to support this study.

## Conflicts of Interest

The authors declare that there is no conflict of interest with any financial organizations regarding the material reported in this manuscript.

## Acknowledgments

This work was supported by the School level youth fund project of Wuhan Donghu University (No. 2019dhs006).

## References

- [1] M. Zhou, X. Li, Y. Wang, S. Li, Y. Ding, and W. Nie, "6G multi-source information fusion based indoor positioning via gaussian kernel density estimation," *IEEE Internet of Things Journal*, vol. 99, pp. 1–1, 2020.
- [2] Z. Ding, S. Liu, L. Liao, and L. Zhang, "A digital construction framework integrating building information modeling and reverse engineering technologies for renovation projects," *Automation in Construction*, vol. 102, pp. 45–58, 2019.
- [3] E. Rytova and S. Gutman, "Assessment of regional development strategy in the context of economy digitization on the basis of fuzzy set method," *IOP Conference Series: Materials Science and Engineering*, vol. 497, no. 1, pp. 12060–12060, 2019.
- [4] A. López, M. Ángeles, and G. P. Beneyto, "Information systems and the computer audit applied to a regional Audit institution: la Sindicatura de Comptes de la Comunidad Valenciana (Spain)," *Revista Gestão e Secretariado-GeSec*, vol. 11, no. 2, pp. 120–138, 2020.
- [5] J. Queiroz, P. Leitão, J. Pontes, A. Chaves, J. Parra, and M. E. Perez-Pons, "A quality innovation strategy for an inter-regional digital innovation hub," *Advances in Distributed Computing and Artificial Intelligence Journal*, vol. 9, no. 4, pp. 31–45, 2020.
- [6] E. Bareinboim and J. Pearl, "Causal inference and the data-fusion problem," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 113, no. 27, pp. 7345–7352, 2016.
- [7] K. Liu and S. Huang, "Integration of data fusion methodology and degradation modeling process to improve prognostics," *IEEE Transactions on Automation Science and Engineering*, vol. 13, no. 1, pp. 344–354, 2016.
- [8] L. Mou, X. Zhu, M. Vakalopoulou et al., "Multitemporal very high resolution from space: outcome of the 2016 IEEE GRSS data fusion contest," *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, vol. 10, no. 8, pp. 3435–3447, 2017.
- [9] B. García de Soto, I. Agustí-Juan, J. Hunhevicz et al., "Productivity of digital fabrication in construction: cost and time analysis of a robotically built wall," *Automation in Construction*, vol. 92, pp. 297–311, 2018.
- [10] Chemical Engineering World, "Hexagon PPM partners with Skanska on HxGN SMART build collaboration accelerates strategic development of digital construction technologies," *Chemical Engineering World*, vol. 52, no. 6, pp. 16–18, 2017.
- [11] E. Vinson, "Applying an established format to the Houston archives bazaar," *Microform & digitization review*, vol. 48, no. 1, pp. 46–53, 2019.
- [12] A. Sudrajat, I. Sudirman, and R. Prasetyo, "Digitalization of logistics processes and comparison with several Asian countries related to logistics information systems: propositions of National Logistics System Architecture," *Solid State Technology*, vol. 63, no. 3, pp. 2824–2836, 2020.
- [13] S. Maksimov, Y. Vasin, N. Valuyskov, and K. Utarov, "The digitization of criminal policy as a tool of overcoming its

- unsystematicity,” *Russian Journal of Criminology*, vol. 13, no. 3, pp. 395–407, 2019.
- [14] I. Ahmad, R. W. Aslam, L. Li et al., “Evaluating focal mechanism of September 24, 2013 Awaran earthquake with geospatial techniques,” *International Journal of Innovations in Engineering and Technology*, vol. 2, no. 3, pp. 108–124, 2021.
- [15] P. B. Fenberg, A. Self, J. R. Stewart, R. J. Wilson, and S. J. Brooks, “Exploring the universal ecological responses to climate change in a univoltine butterfly,” *Journal of Animal Ecology*, vol. 85, no. 3, pp. 739–748, 2016.
- [16] W. Liao, X. Huang, F. Van Coillie et al., “Processing of multi-resolution thermal hyperspectral and digital color data: outcome of the 2014 IEEE GRSS data fusion contest,” *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, vol. 8, no. 6, pp. 2984–2996, 2015.
- [17] M. Berkow, C. M. Monsere, P. Koonce, R. L. Bertini, and M. Wolfe, “Prototype for data fusion using stationary and mobile data: sources for improved arterial performance measurement,” *Transportation Research Record*, vol. 2099, no. 1, pp. 102–112, 2018.
- [18] A. Amamra and N. Aouf, “Real-time multiview data fusion for object tracking with RGBD sensors,” *Robotica*, vol. 34, no. 8, pp. 1855–1879, 2016.
- [19] N. Yokoya, C. Grohnfeldt, and J. Chanussot, “Hyperspectral and multispectral data fusion: a comparative review of the recent literature,” *IEEE Geoscience & Remote Sensing Magazine*, vol. 5, no. 2, pp. 29–56, 2017.
- [20] R. Cabral Farias, J. E. Cohen, and P. Comon, “Exploring multimodal data fusion through joint decompositions with flexible couplings,” *IEEE Transactions on Signal Processing*, vol. 64, no. 18, pp. 4830–4844, 2016.
- [21] K. M. Nunes, M. V. O. Andrade, A. M. P. Santos Filho, M. C. Lasmar, and M. M. Sena, “Detection and characterisation of frauds in bovine meat *\_in natura\_* by non- meat ingredient additions using data fusion of chemical parameters and ATR- FTIR spectroscopy,” *Food Chemistry*, vol. 205, pp. 14–22, 2016.
- [22] F. Xiao, “Multi-sensor data fusion based on the belief divergence measure of evidences and the belief entropy,” *Information Fusion*, vol. 46, no. 46, pp. 23–32, 2019.
- [23] W. Sun, X. Zhang, Z. Zhang, and R. Zhu, “Data fusion of near-infrared and mid-infrared spectra for identification of rhu-barb,” *Spectrochimica Acta Part A Molecular & Biomolecular Spectroscopy*, vol. 171, pp. 72–79, 2017.
- [24] A. Majumder, L. Behera, and V. K. Subramanian, “Automatic facial expression recognition system using deep network-based data fusion,” *IEEE Transactions on Cybernetics*, vol. 48, no. 99, pp. 103–114, 2017.
- [25] D. Zhou, K. Zhang, A. Ravey, F. Gao, and A. Miraoui, “On-line estimation of lithium polymer batteries state-of-charge using particle filter based data fusion with multi-models approach,” *IEEE Transactions on Industry Applications*, vol. 52, no. 3, pp. 2582–2595, 2016.
- [26] F. C. Chen, M. R. Jahanshahi, R. T. Wu, and C. Joffe, “A texture-based video processing methodology using Bayesian data fusion for autonomous crack detection on metallic surfaces,” *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 4, pp. 271–287, 2017.
- [27] Z. Tao and P. Bonnifait, “Sequential data fusion of GNSS pseudoranges and Dopplers with map-based vision systems,” *IEEE Transactions on Intelligent Vehicles*, vol. 1, no. 3, pp. 254–265, 2016.
- [28] M. Zhou, Y. X. Long, W. P. Zhang et al., “Adaptive genetic algorithm-aided neural network with channel state information tensor decomposition for indoor localization,” *IEEE Transactions on Evolutionary Computation*, vol. 25, no. 5, pp. 913–927, 2021.