Open Access Library Journal
2025, Volume 12, e13225
ISSN Online: 2333-9721
ISSN Print: 2333-9705

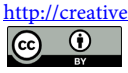
Research on the Design of Data Asset Valuation Platform Based on WEB

Tian Shen

School of Computer Science and Technology, Zhejiang Normal University, Jinhua, China
Email: 799607132@qq.com

How to cite this paper: Shen, T. (2025) Research on the Design of Data Asset Valuation Platform Based on WEB. *Open Access Library Journal*, **12**: e13225.
https://doi.org/10.4236/oalib.1113225

Received: March 7, 2025
Accepted: April 6, 2025
Published: April 9, 2025



Abstract

With the rapid development of big data, mining the potential value of big data has become the key to the future development of science and technology. Especially with the introduction of the concept of "data assets", popularizing data asset knowledge and expanding the data asset market have gradually become research hotspots. As an intangible asset, data assets have the characteristics of continuous growth, diverse categories, and frequent updates compared to traditional assets. Therefore, it is particularly important to build a Web-based data asset platform equipped with a complete backend management system. The research aims to design and develop a data asset valuation platform that can run on the web page. The development steps include: collecting and crawling data asset information on the JD Wanxiang platform, and cleaning the data; based on the cleaned data, providing users with data asset valuation, personalized recommendations, market situation analysis and other functions on the front end. The main functions of the backend management system include user management, data asset information management, and daily market data management.

Subject Areas

Physical Chemistry

Keywords

Data Assets, Intelligent Recommendation, Content-Based Recommendation, Backend Management System

1. Introduction

In the information age, the status of big data has been comprehensively improved and has become a hot word in recent years. Advertisers hope to reach profitable

consumers through data, medical experts hope to find the side effects of prescription drugs, supply chain operators hope to optimize their delivery routes, police hope to determine where to concentrate resources, and social scientists hope to study human interactions [1]. We can see that in the Internet age, big data technology is increasingly becoming a huge driving force for industry transformation and development and maintaining vitality [2].

The focus of government data management from the perspective of assets is the quantification of data value and the realization of data value [3]. From the perspective of technological progress and productivity development, big data belongs to technical resources. With the rapid development of science and technology, technical resources are playing an increasingly decisive role in economic and social development. Big data is being transformed from resources into assets and then into funds. Not only is the value of data increasing day by day, but it has also become a key factor in enhancing the core competitiveness of enterprises [4]. How to define data assets and evaluate the value of data assets has become a topic of increasing importance to the government and enterprises. How to obtain relevant information about data assets more conveniently and quickly has become a major issue that needs to be solved urgently.

2. Significance

In today's era of "everything is data", massive amounts of data are generated every minute and every second. These data assets are not only huge in quantity, but also have huge commercial value. Take LinkedIn as an example. This is a world-renowned workplace social networking site company. When it was acquired, its user base had reached an astonishing 400 million members, and its annual revenue exceeded US$3 billion [5] [6]. This example fully demonstrates the huge profits that data assets can bring to governments and enterprises. Therefore, a data asset platform that can meet the needs of governments and people engaged in data assets will surely become one of the main ways for all walks of life to understand data assets in the future. Such a platform will become an important means to reduce valuation costs, increase data acquisition methods, and increase the value of data assets.

In order to further promote the development of the data market, we are researching and developing a platform that integrates data asset valuation and personalized recommendation functions. This platform can not only provide users with daily sales information of data assets, but also facilitate users to understand the market situation and make wise decisions while gaining a deep understanding of their own data assets. Through such a platform, users will be able to participate in the data market more effectively and promote the development of the data market.

3. Analysis of Research Status

3.1. Data Asset Valuation Algorithm Related

We categorized data assets into three types based on transaction types: API, data

package, and data report. Furthermore, we refined the features across six major dimensions: temporal characteristics, overall market operation characteristics, enterprise information characteristics, individual operation characteristics, data asset attributes, and consumer behavior characteristics. After feature merging, we ultimately obtained 20 labels for the API category and 23 labels for both data packages and data reports. The project selected five ensemble learning models—XGBoost, LightGBM, AdaBoost, GBR, and Random Forest—for data asset valuation. Under the default hyperparameter settings, XGBoost and LightGBM demonstrated superior performance. Therefore, hyperparameter tuning was further applied to these two models. Ultimately, LightGBM achieved the best performance after optimization, with an RMSLE of 0.079, outperforming XGBoost's 0.102 (See Table 1 and Table 2).

Table 1. Data asset API characterization table.

Data Asset Category	Feature Type or Construction Rule
Data Asset Entry Name	Raw Feature
Data Asset Views	Raw Feature
Data Asset Likes	Raw Feature
Data Asset Favorites	Raw Feature
Data Asset Transactions	Raw Feature
Number of Entries in Primary Data Asset Category	Count of API products in each primary category
Number of Entries in Secondary Data Asset Category	Count of API products in each secondary category
Data Asset Transaction Volume	Raw Feature
Total Transaction Price of Data Asset	Raw Feature
Data Provider Registration Year	Raw Feature
Data Provider Company Type	Raw Feature
Data Provider Company Status	Raw Feature
Data Provider Registered Address	Raw Feature
Data Provider Registered Capital	Raw Feature
Number of Products on the Platform by Data Provider	Number of data asset products owned by each provider, calculated based on merchant ID
Data Asset Transactions in the Past Six Months	Transaction volume calculated from transaction records
Data Asset Transactions in the Past Year	Transaction volume calculated from transaction records
Data Asset Repurchase Count in the Past Six Months	Number of times users repurchased, based on transaction records
Data Asset Repurchase Count in the Past Year	Number of times users repurchased, based on transaction records
Data Asset Transaction Unit Price	Total transaction price/transaction volume

Table 2. Data asset data package and data report characterization table.

Data Asset Category	Feature Type or Construction Rule
Data Asset Entry Name	Raw Feature
Data Asset Views	Raw Feature
Data Asset Likes	Raw Feature
Data Asset Favorites	Raw Feature
Data Asset Transactions	Raw Feature
Number of Entries in Primary Data Asset Category	Count of data package and data report products in each primary category
Number of Entries in Secondary Data Asset Category	Count of data package and data report products in each secondary category
Data Asset File Size	Raw Feature
Data Asset File Type	Raw Feature
Date Associated with Data Asset File	Raw Feature
Year Associated with Data Asset File	Extracted from the date
Whether the Data Asset File is Time-Sensitive	Determine whether time information is missing in the product description
Data Provider Registration Year	Raw Feature
Data Provider Company Type	Raw Feature
Data Provider Company Status	Raw Feature
Data Provider Registered Address	Raw Feature
Data Provider Registered Capital	Raw Feature
Number of Products on the Platform by Data Provider	Number of data asset products owned by each provider, calculated based on merchant ID
Data Asset Transaction Unit Price	Total transaction price/transaction volume
Data Asset Category	Feature Type or Construction Rule

3.2. Recommendation Algorithm Related

In today's era of information explosion, the application of recommendation algorithms has become ubiquitous. From search engines to online learning platforms to various shopping applications, recommendation systems play a vital role. They provide users with personalized recommendation services by analyzing data such as users' browsing history, purchasing behavior, and search habits. In recent years, researchers at home and abroad have conducted extensive and in-depth research on recommendation algorithms and their service models. For example, Hwang and his colleagues proposed a new recommendation algorithm to improve the accuracy of recommendations by focusing on the mining of movie genres in order to improve the accuracy of recommendations in response to the shortcomings of traditional collaborative filtering recommendation algorithms in terms of accuracy [7]. Phonexay *et al.* proposed a recommendation algorithm based on an im-

proved k-set, which can effectively establish analysis groups in social networks to better understand the relationships and interests between users [8]. Shu *et al.* proposed a method in their research report that uses convolutional neural networks (CNNs) to extract text information from items. By analyzing the latent elements of items, the method predicts the user's rating of the items and uses this as an indicator of the user's recommendation rating, thereby recommending products or content that users may be interested in [9].

In the current market, mainstream recommendation algorithms mainly include collaborative filtering recommendation, content-based recommendation, popularity-based recommendation, and association rule-based recommendation. Each of these recommendation algorithms has its own unique advantages and limitations. Collaborative filtering algorithms can be divided into two types: item-based and user-based. The item-based collaborative filtering algorithm recommends similar items based on the items that the user liked in the past; while the user-based collaborative filtering algorithm recommends items to the user based on items liked by other users with similar interests to the user. However, a major drawback of this algorithm is that it cannot solve the cold start problem of items, that is, newly added items are difficult to be recommended to users due to the lack of sufficient user data. Content-based recommendation algorithms push items with similar content based on the content features of items that users have liked or paid attention to in the past. A significant advantage of this algorithm is that it can avoid the cold start problem of items and is relatively simple to implement [10].

3.3. Data Asset Related

At present, as a key production factor of the digital economy, data is growing and accumulating at an unprecedented exponential rate, becoming a type of resource [11] [12]. The concept of data assets is derived from the concepts of information resources and data resources. It was born in the 1970s. When it was first mentioned, its main content was assets such as government bonds, corporate bonds and physical bonds that could be held. At the World Economic Forum in 2011, it was first proposed that data assets be included in the asset category. As an intangible asset, data assets have many unique characteristics, such as information value, such as frequent updates and diverse categories; value-added attributes that continuously create new value; and unlimited sharing at zero cost.

There are many difficulties in the process of data assetization. In addition to the fact that data is easily leaked due to its characteristics, it is worth discussing who is the final owner of the data during the data transaction process.

3.4. Flask Framework Introduction

Flask is a lightweight web application framework that uses Python. Compared with backend development frameworks, Flask is more flexible, lightweight and easy to use [11]. It can be well combined with the MVC development model, al-

lowing developers to implement small and medium-sized websites with rich functions in a short period of time. In addition, the Flask framework is also highly customizable. Developers can add corresponding functions according to their needs, while keeping its core functions simple and enriching the functions implemented. The plug-in library in Flask provides developers with a variety of functions, allowing users to achieve personalized customization during use and develop powerful web applications.

3.5. Introduction to Vue.js Framework

Vue is a progressive framework used to build user interaction interfaces [12]. The design concept of Vue is to apply layer by layer from the bottom up. Its objects separate the view and the data, so that when developers use this framework to program, they only need to focus on the view layer and do not need to confirm the DOM object. It is not only easy to use, but also can be easily integrated with other libraries, such as Vant-UI. In addition, when Vue is used in combination with various supporting libraries, it can also provide good services for complex single-page applications.

3.6. Introduction to jQuery.js Framework

jQuery is a simple and fast framework based on JavaScript. It encapsulates many commonly used functional codes in JavaScript, making the code more concise during the development process [3]. At the same time, this framework also optimizes the Ajax operation function used when interacting with the back end, so that Web projects do not need to refresh the entire web page when data is updated, but only need to partially refresh the part that uses the data (see **Figure 1**).

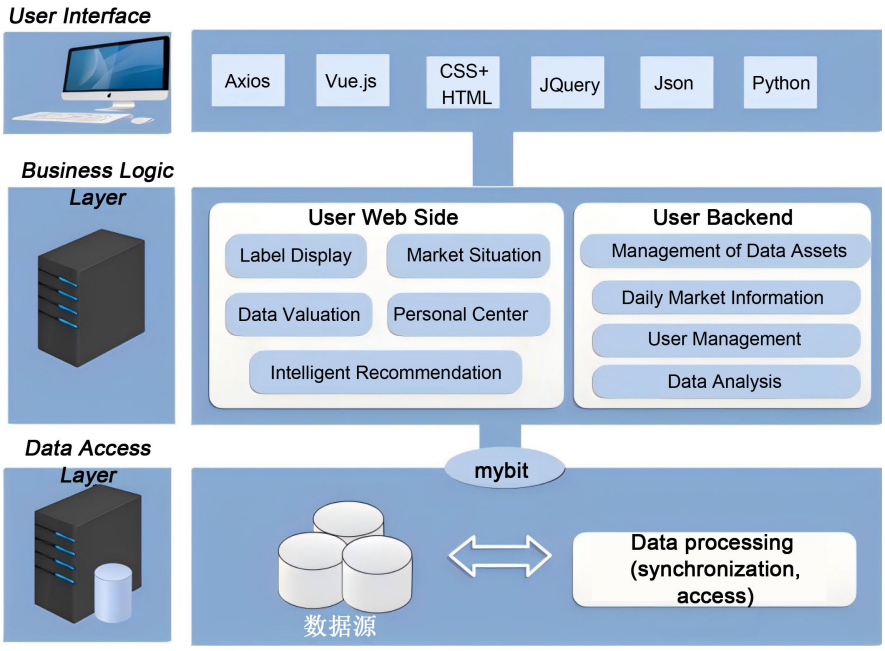


Figure 1. System three-layer architecture diagram.

4. System Overview Design

In the system outline design stage, the research mainly focuses on the overall architecture of the data asset valuation platform and the design of its various components. The platform will adopt a layered architecture design model to ensure clear responsibilities between each layer, which is convenient for subsequent system maintenance and functional expansion. The system architecture will be subdivided into the front-end display layer, business logic layer and data storage layer, and each layer will have specific functions and responsibilities. The front-end display layer will be responsible for interacting with users and providing a user-friendly interface; the business logic layer will be responsible for processing various business rules and processes; and the data storage layer will focus on data storage and management. Starting with the explanation of the global architecture of the system, this study designs corresponding functional modules according to the needs of different end users, and gives a detailed introduction to the database table structure.

4.1. System Global Architecture

The system adopts the currently popular MVC three-tier architecture model, namely model, view, and controller. This model realizes the separation of the presentation layer and the business logic layer, thereby improving the maintainability, portability, scalability, and reusability of the program, while reducing the complexity of program development and ensuring the loose coupling characteristics of the system.

4.2. System Module Design

User Web Terminal Function Module Design

1) Hierarchical label display module

Industry classification label display: Data assets are classified and displayed according to industry, covering secondary industry label information under five major categories to enhance users' clear understanding of data asset classification.

Application classification label display: Data assets are displayed based on three application classifications: API, data package, and data report. This deepens users' understanding of application and industry classification and their relationship, and assists users in utilizing data asset valuation functions.

2) Data market situation module

Overall market situation: Provide users with data market daily and weekly sales rankings, total sales rankings and transaction amount rankings, and update and display the top data of each ranking every day, so that users can understand the transaction dynamics of the data asset market and reasonably improve the value of data assets.

Situation of each application industry: It shows users the transaction and sales market situation of three application types, API, data package and data report, in the form of a line graph. Users can observe the market transaction trend through

the changes in the line graph.

3) Data intelligent valuation module

Input module: Users can input the data assets that need to be valued through the input module. The system provides single data input and modular partial input options. The tags related to the data assets entered by the user will be stored so that the user can directly select them the next time they enter.

Valuation module: Users select a valuation algorithm for valuation, and the platform will provide recommended results. Users can rate and comment on the data valuation results and submit feedback to the system backend to provide a reference for administrators to optimize the data valuation algorithm.

Generate data report: Users can choose to print data reports and save the valuation data assets and their results in PDF format locally.

4) Data intelligent personalized recommendation module

The platform provides a personalized recommendation function, which recommends data assets that users may be interested in based on their historical valuation data and collected data assets. Users can access relevant data asset information in the smart recommendation section to further understand data assets of the same type.

5) Personal Center Module

Personal information modification: The platform provides the function of viewing and modifying personal information. Users can modify personal information including user name, mobile phone number and email address in this section.

Historical data list and collection: This function shows users the data assets that have been historically valued. Users can view past data reports and collected data asset content.

6) Design of functional modules of backend management

Login: The backend administrator enters the backend management system through the login module and uses the corresponding management functions.

User management: Administrators manage and maintain user information through the user management module, including management of user account creation, information modification, account deletion, user information, data asset collection and valuation information.

Data asset information management: Administrators use this module to manage and maintain data asset information and assist in data analysis.

Data market daily information management: Administrators can use this module to view the daily data market information crawling status in order to update the crawler code.

4.3. Design and Implementation of Backend Management System

1) Overall design ideas of the background

The business process of this user backend management system is as follows: the administrator enters the system backend homepage from the login interface and

selects the required functional modules through the sidebar, including the homepage, user management, data asset management, and daily data management modules. In addition, it also includes a crawler module for daily information acquisition. First log in, after the administrator enters the account password, he successfully logs in to the backend management homepage; then the data asset module, you can view the current three types of data asset content, and query and edit the required data assets; the daily data management module can view the market data crawled by the crawler on that day, so as to modify the crawler code to maintain system operation when encountering anti-crawling or crawling website structure changes; the user management module can view all user information and perform management operations.

2) Specific implementation of user backend management system

The backend management system is built with Vue framework and element UI component library, and adopts component development mode. First, the login and registration module is implemented, and the Element form is used for format verification. The login and registration interface is written in the background Flask to verify the login personnel's permissions and realize the login and registration function. The data asset information management module can view all the data assets currently crawled, and perform retrieval and editing operations. The data display uses the Pagination component in the Element component library to set the amount of data displayed per page, and display the total amount of data in the current database, which is convenient for analyzing the fluctuation of the commodity information of the crawling platform. The administrator can choose to edit the data asset information. After clicking the edit button, the data is modified in the editing interface to complete the data asset information editing. The market daily information management module displays the market data obtained by the daily crawler. Since the crawling object may change, such as the change of the web page structure, the administrator can use this module to determine whether the crawler program has an error, and can retrieve information on this page. The user management module can view the platform registered user information, and perform query, edit and delete operations (See **Figure 2**).

5. Summary and Outlook

5.1. Summary

The research and design can realize a data asset valuation system based on WEB. Through the joint use of Flask framework and Vue.js framework, a data asset valuation platform with complete functions and user-friendly interface can be successfully built. The platform includes multiple modules such as data asset information management, market daily information management, and user management. Each module realizes the corresponding functional requirements.

In the data asset information management module, users can clearly see all the data assets currently crawled, and can perform operations such as retrieval and editing, which greatly improves the convenience of data management. The market

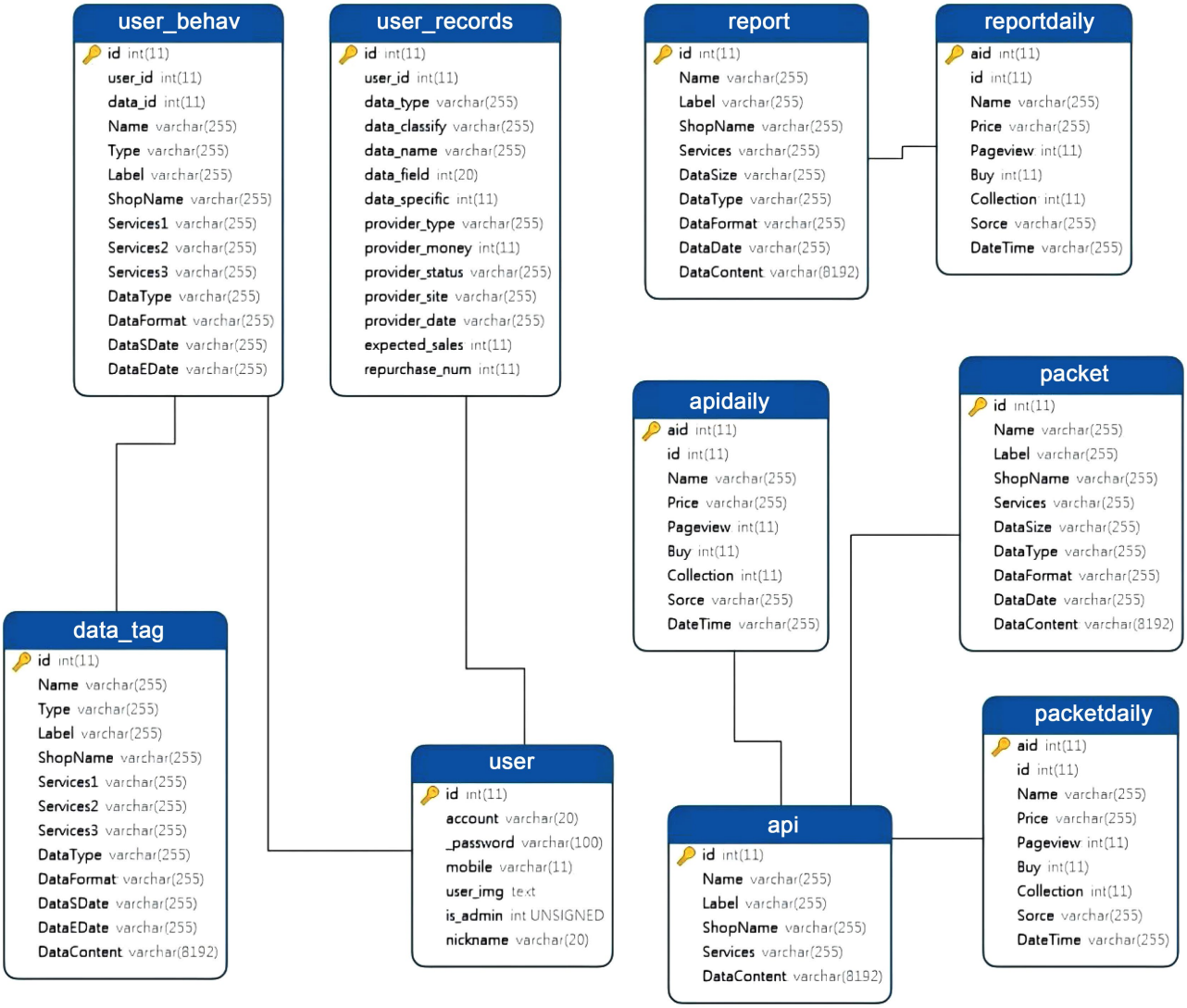


Figure 2. Backend database relationship diagram.

daily information management module can display the market data obtained by the daily crawler, helping administrators to determine whether there is an error in the crawler program, and also supports the retrieval function of information. The user management module provides comprehensive management of user information, including operations such as query, edit and delete, ensuring the security and accuracy of user information. In the system design, user needs are fully considered. Through modular design ideas, multiple core functions such as hierarchical label display, data market analysis, data intelligent valuation, data personalized recommendation and personal center are realized, aiming to provide users with comprehensive and convenient data asset management services. At the same time, the development of the background management system also fully considers practicality and efficiency. Through the construction of the user management module and the distributed crawler system, effective management of user data and efficient execution of data capture are realized.

5.2. Outlook

With the continuous development of big data technology, the data asset valuation platform will play a more important role in the future. In future research, we will further optimize the performance of the system and improve the efficiency and accuracy of data processing. At the same time, we will continue to explore new algorithms and technologies to better meet users' needs for data asset valuation. In addition, we will also strengthen the security and stability of the system to ensure the security and privacy of user data. I believe that with our continuous efforts, the data asset valuation platform will provide more users with better quality and more convenient services, and promote the continuous development and application of big data technology. In the future, we will continue to deepen our research in the field of data assets, explore more innovative application scenarios, and strive to develop a more intelligent and efficient data asset management platform to contribute to the development of the data economy.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] Barocas, S. and Selbst, A.D. (2016) Big Data's Disparate Impact. *California Law Review*, **104**, 671-732. https://doi.org/10.2139/ssrn.2477899

[2] Tang, S.Y. (2020) Research on the Development Model and Current Status of Big Data + Asset Evaluation. *China Business Review*, No. 13, 43-45.

[3] Zhu, D. (2017) Research on the Value Assessment and Realization of Government Data Assets. South China University of Technology.

[4] Li, M.Y. (2020) On Transforming Big Data Resources into Big Data Assets.

[5] Baidu Wenku (2019) Report on the Road to Data Assetization: Valuation and Industry Practice of Data Assets. https://wenku.baidu.com/view/709ccae6f08583d049649b6648d7c1c709a10bf3.html?_wkts_=1743590612567&bdQuery=%E6%95%B0%E6%8D%AE%E8%B5%84%E4%BA%A7%E5%8C%96%E4%B9%8B%E8%B7%AF--%E6%95%B0%E6%8D%AE%E8%B5%84%E4%BA%A7%E7%9A%84%E4%BC%B0%E5%80%BC%E4%B8%8E%E8%A1%8C%E4%B8%9A%E5%AE%9E%E8%B7%B5%E3%80%8B%E6%8A%A5%E5%91%8A

[6] Huang, D.X. and Pan, X.F. (2022) Research and Practice on Data Asset Value Calculation. *Information and Communication Technology and Policy*, No. 2, 29-36.

[7] Hwang, T., Park, C., Hong, J. and Kim, S.K. (2016) An Algorithm for Movie Classification and Recommendation Using Genre Correlation. *Multimedia Tools and Applications*, **75**, 12843-12858. https://doi.org/10.1007/s11042-016-3526-8

[8] Vilakone, P., Park, D., Xinchang, K. and Hao, F. (2018) An Efficient Movie Recommendation Algorithm Based on Improved K-Clique. *Human-Centric Computing and Information Sciences*, **8**, 1-15. https://doi.org/10.1186/s13673-018-0161-6

[9] Shu, J., Shen, X., Liu, H., Yi, B. and Zhang, Z. (2017) A Content-Based Recommendation Algorithm for Learning Resources. *Multimedia Systems*, **24**, 163-173. https://doi.org/10.1007/s00530-017-0539-8

[10] Qin, C., Zhao, T.Z. and Liu, Y. (2021) A Review of the Research and Development of Personalized Recommendation Algorithms. *Journal of Dongguan University of Technology*, **28**, 51-60.

[11] Shen, H.W. (2018) Design and Implementation of Enterprise Intranet Security System Based on Flask. Beijing Jiaotong University.

[12] Chen, X., Li, J. and Qiu, Q. (2020) Smart Traffic Program Solution Based on Vue.js Framework. *Information Technology and Informatization*, No. 11, 240-243.