Research on Data Asset Assessment Methodology and its Application in the Transformation of New Quality Productivity

Weili Li

School of Management, China University of Mining and Technology (Beijing), Beijing 100083, China

15201308486@163.com

Abstract. As data emerges as a critical production factor, research on data asset valuation and its role in productivity transformation has gained significant attention in both academic and business circles. This paper systematically reviews the existing data asset evaluation methods, including the cost, market, and income approaches, discussing their respective strengths, limitations, and suitable application contexts. Building on this foundation, we propose a deep learning-based data asset valuation model that incorporates multi-dimensional feature extraction, dynamic weight allocation, and an error adjustment mechanism, enabling more precise estimation of data assets' economic potential. Additionally, we develop a model for converting data assets into new productive forces, examining their value-added impact in real-world production and quantifying their contribution to productivity using multi-level regression analysis and deep neural networks. Experimental results reveal a strong correlation between data assets' value-added outcomes and their transformation efficiency into new productive forces, with dataset characteristics significantly influencing the transformation results.

Keywords: Data asset valuation, new quality productivity transformation, deep learning, data trading models

1 Introduction

As data becomes a vital resource in the modern economy, its valuation and management are increasingly emphasized by businesses and researchers[1]. Unlike tangible assets, the value of data includes its potential to enhance decision-making, market analysis, and innovation. Traditional valuation methods fall short in capturing the complex value of data assets, highlighting the need for more accurate approaches[2]. Data assets not only store and transmit information but also act as production factors that generate economic benefits. In the digital economy, effective use of data assets drives productivity, technological innovation, resource optimization, and market expansion, making the shift from static resources to dynamic productivity drivers a key focus of data economy research.

K. Zhang et al. (eds.), *Proceedings of the 5th International Conference on Economic Management and Big Data Application (ICEMBDA 2024)*, Advances in Economics, Business and Management Research 313,
https://doi.org/10.2991/978-94-6463-638-3_51

This paper introduces a deep learning-based model for data asset valuation that leverages the strengths of traditional methods like the market and income approaches. It aims to facilitate the conversion of data assets into new-quality productivity across different application scenarios. Through empirical research and experimental analysis, this study not only addresses theoretical gaps in current data asset valuation methods but also offers practical guidance for enterprises seeking to optimize data utilization in digital transformation efforts.

2 Related work

2.1 Data Asset Valuation Methodology

Traditionally, the cost approach to valuing data assets focused on expenses related to data generation and management. Recent research has expanded this approach to include full life-cycle costs, hidden costs (e.g., data cleansing), and long-term storage expenses, providing a more comprehensive valuation. Wang and Zhao (2020)[3] incorporated both domestic and international perspectives on data asset valuation. TLi et al. (2024)[4] developed an indicator system to assess data asset value based on quality, application, and risk. Additionally, dynamic earnings models have been introduced to adjust valuations over time. Kim et al. (2021)[5] examined the income model, identifying key elements and considerations for valuing data held by companies.

2.2 Study on Data Asset Trading Model

The on-market trading model for data assets operates on centralized platforms, similar to traditional financial markets, using standardized processes, pricing mechanisms, and contract templates to enhance transparency and security. Over-the-counter (OTC) trading uses centralized order books or bidding to aggregate trades and provide prompt responses. This model, common in financial markets, is now being adapted for data assets. Deng and Liu (2024) [6] proposed a pricing mechanism combining market-based and government intervention for legal compliance. As data assets grow in importance, legal frameworks are emerging to protect stakeholders' rights. Within this model, data assets can be securitized, increasing liquidity and attracting investors.

3 Model

3.1 Data Asset Valorization Valuation Model

To accurately assess data value within complex, multidimensional big data environments, this paper introduces a data asset valorization valuation model. The data asset valorization assessment process is illustrated in Fig. 1.

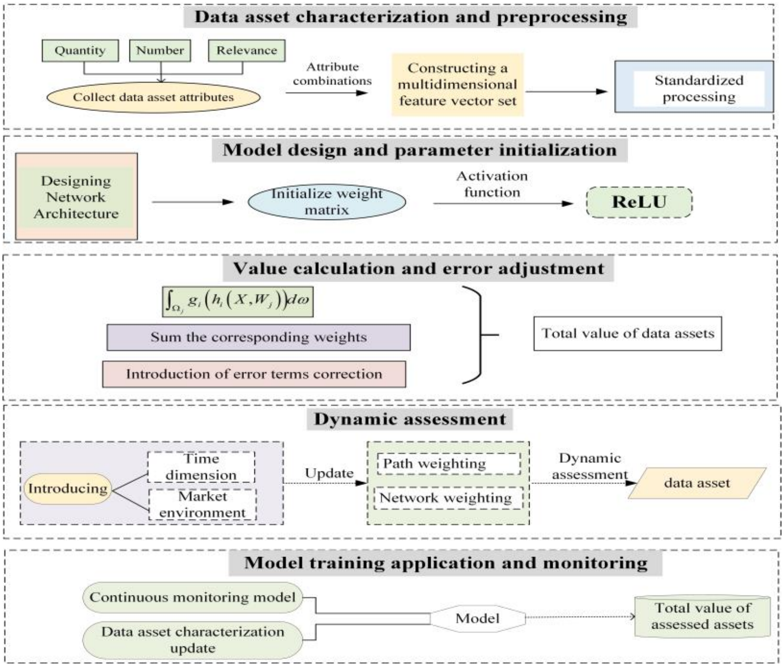

Fig. 1. Flowchart for valuing data assets

Let the total value of the data asset be V_d, which is calculated by the formula

$$V_d = \sum_{j=1}^{k} \gamma_j \cdot \left(\int_{\Omega_j} g_j\left(h_j(X, W_j)\right) d\omega\right) + \epsilon, \tag{1}$$

where X is a multidimensional set of feature vectors representing attributes of the data asset, such as quality (Q), quantity (N), and relevance (R). The feature vectors may be further represented as $X = \{Q, N, R, \ldots\}$, $h_j(X, W_j)$ is a deep neural network transform function of the j path for extracting high-dimensional nonlinear features. W_j is a weight matrix on the path. $g_j(\cdot)$ is the activation function of the output layer. Ω_j denotes the feature space over the paths, and $d\omega$ is the integral element. γ_j is the weight parameter of the j path. ϵ is the error term.

In addition, considering the dynamic nature of the data asset valuation process, the time dimension and the influence of the market environment are introduced:

$$V_d(t) = \sum_{j=1}^{k} \gamma_j\,(t) \cdot \left(\int_{\Omega_j} g_j\left(h_j\left(X(t), W_j(t)\right)\right) d\omega\right) + \epsilon(t) \tag{2}$$

where t denotes the time dimension. $\gamma_j(t)$ and $W_j(t)$ are the path weight parameter and network weight matrix under time t, respectively, taking into account the dynamic adjustment of the model by changes in the market environment.

Advanced adaptive learning rate tuning Adam algorithms are used. The optimizer is able to adjust the learning rate in real time according to the training state of the model, improving the model's adaptability in different market environments.

3.2 Data Asset Effective Transformation Model

The data asset effective transformation model integrates multi-level regression and deep neural networks to quantify productivity gains from data assets, supporting enterprise decision-making (see Fig. 2).

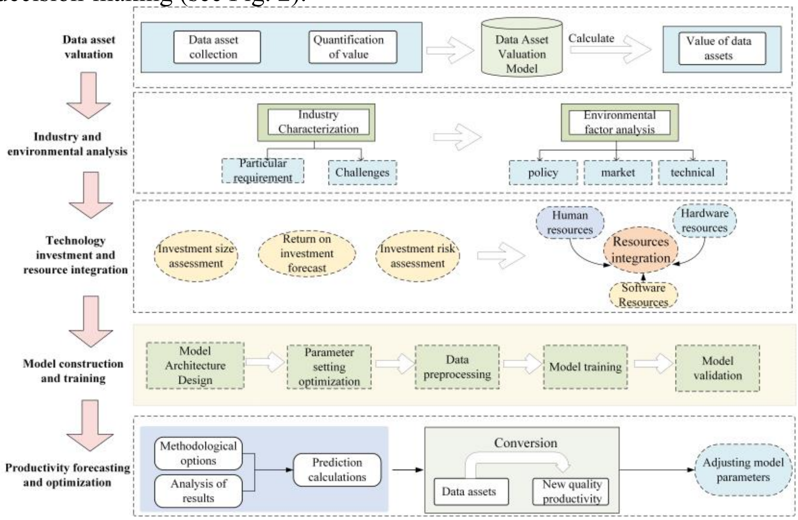

Fig. 2. Flowchart of the model for effective transformation of data assets into new quality productivity

The true value of data assets lies in their ability to generate value-added benefits during the production process. By integrating deep learning with multi-level regression analysis, the proposed model more accurately quantifies the contribution of data assets to productivity, providing guidance for enterprises to better leverage data resources and enhance economic efficiency. Let P_f represent the new quality productivity generated by data assets, calculated as follows:

$$P_f = \alpha \cdot \left(\sum_{i=1}^{n} \left(V_d^{(i)} \cdot T^{(i)}(X, Z) \right)^{\beta_1} \cdot E_i \cdot G\left(\frac{I_i(t, Z)}{L_i(X)} \right) \right)^{\beta_2} + \delta \cdot H(Y, t) + \eta, \tag{3}$$

where α is the baseline productivity factor, which represents the underlying contribution of the data asset to productivity. $V_d^{(i)}$ is the value of the data asset of type i. $T^{(i)}(X, Z)$ is the transformation coefficient of the data asset of type i, which depends on specific data characteristics X and industry environment factors Z. β_1 and β_2 are weighting parameters. E_i is the utilization efficiency of the data assets of the i category. $G(\cdot)$ is the augmentation function of the impact of external technology and resources on productivity, indicating the utility of technology investment $I_i(t, Z)$ relative to input resources $L_i(X)$. δ is the coefficient of external effects. $H(Y, t)$ is a function of the impact of external investment Y on productivity over time t. η is the error term.

The conversion factor $T^{(i)}(X, Z)$ can be further refined as:

$$T^{(i)}(X,Z) = \frac{I^{(i)}(t,Z)\cdot E^{(i)}(X)}{L^{(i)}(X)\cdot R^{(i)}(Z)\cdot \Gamma(Z,t)} \tag{4}$$

where $I^{(i)}(t, Z)$ is the technological investment in the time t and environment Z of the class i data asset. $E^{(i)}(X)$ is the integration and utilization efficiency. $L^{(i)}(X)$ is the labor or resource investment required. $R^{(i)}(Z)$ is the transformation rate of the Class I data asset in the industry-specific environment, and $\Gamma(Z, t)$ is a correction function of the environmental and temporal factors.

4 Experimentation and Analysis

4.1 Experimental Datasets

(1) Shanghai Residential Investment Data: This dataset from the Shanghai Statistics Bureau contains residential investment figures and their percentage share of total fixed asset investment.

(2) Guangzhou Market Price Data: Sourced from the Guangzhou Public Data Open Platform, this dataset offers insights into market conditions, including data ID, product type, name, market, specifications, average price, unit, date, and update time.

(3) Guangzhou Industrial Value-Added Data: Released by the Guangzhou Municipal Bureau of Statistics, this dataset provides value added by industries of a certain scale, featuring current and cumulative data, dates, year-on-year growth percentages.

4.2 Experimental Setup

The experiments were conducted using PyTorch 2.1.2 on a Windows system with an Nvidia RTX A4000 GPU and an Intel Xeon W-2175 CPU. A deep neural network with three fully connected layers (128, 64, and 32 neurons) and ReLU activation was trained for 1000 epochs, with loss values reported every 100 epochs. We use 5-fold cross-validation to validate the model. Through multiple training and validation, it makes the evaluation results more stable and reliable, and can effectively reduce the bias caused by uneven data distribution.

4.3 Data Asset Assessment Results and Analysis

The assessment results, shown in Fig. 3, indicate that the Guangzhou industrial value-added dataset has the highest data value (5.392), thanks to its extensive, diverse time-series information across multiple variables, which supports multidimensional analysis and forecasting. In contrast, the Shanghai residential investment dataset (2.848) and the Guangzhou market price dataset (0.747) have lower values due to limited time-series coverage. The Shanghai dataset includes only three time-varying columns, while the Guangzhou market price dataset covers a shorter timeframe, limiting their analytical capabilities.

Data richness and representativeness also contribute significantly to data value. The Guangzhou industrial value-added dataset excels in these areas, offering broad coverage across various industries and dimensions, making it suitable for multiple applications. In contrast, the Shanghai residential investment dataset focuses solely on residential investment, limiting its diversity. The Guangzhou market price dataset, primarily offering price data, shows the lowest richness and weakest representativeness, reflecting only short-term price trends. Consequently, the Guangzhou industrial value-added dataset is the most valuable, followed by Shanghai's residential investment dataset, with the Guangzhou market price dataset ranked lowest.

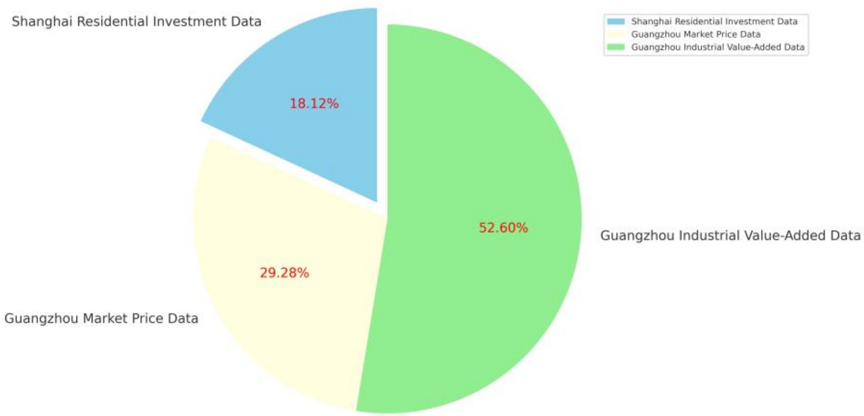

Fig. 3. Comparison of data asset value assessment results for different datasets

4.4 Results and Analysis of Effective Productivity Transformation of Data Assets

The quantitative experiment on the transformation of data assets into new quality productivity results, displayed in Figure 4. The comparison of transformation results and data asset values across the three datasets reveals a strong correlation between the two. Guangzhou's above-scale industrial value-added dataset, with the highest transformation result (36.48) and data value (5.392), demonstrates that its rich time series and multidimensional indicators effectively support complex analysis, modeling, and forecasting, leading to better productivity transformation. The Shanghai residential investment dataset ranks in the middle, with a transformation result of 12.57 and a data value of 2.848, benefiting from some time-series information but lacking the diversity of the industrial dataset. In contrast, the Guangzhou market price dataset has the lowest transformation result (12.84) and data value (0.747), as its single-column, short-span price data limits its analytical depth and application, resulting in the least effective productivity transformation.

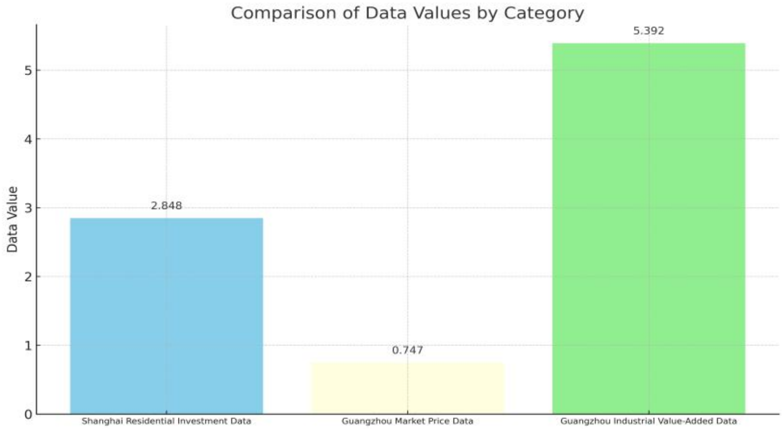

Fig. 4. Comparison of the results of effective new mass productivity transformation

5 Conclusion

This paper systematically examines data asset valuation methods and their role in transforming new quality productivity, introducing a deep learning-based valuation model and verifying its effectiveness across various scenarios. The findings emphasize that accurately assessing the multidimensional characteristics of data assets is crucial for productivity enhancement. Future research should focus on optimizing the model's dynamic adaptability and exploring transformation efficiency in complex scenarios. Additionally, integrating blockchain and other technologies could enhance transaction transparency and security.

References

1. Piryonesi S M, El-Diraby T E. Data analytics in asset management: Cost-effective prediction of the pavement condition index[J]. Journal of Infrastructure Systems, 2020, 26(1): 04019036.
2. Birch K, Cochrane D T, Ward C. Data as asset? The measurement, governance, and valuation of digital personal data by Big Tech[J]. Big Data & Society, 2021, 8(1): 20539517211017308.
3. Yanlin W, Haijun Z. Data asset value assessment literature review and prospect[C]//Journal of Physics: Conference Series. IOP Publishing, 2020, 1550(3): 032133.
4. Li A, Wang A, Chi Y, et al. Exploration of Data Asset Valuation for Internet Platform Companies Based on the DEA-BCC Model[J]. Procedia Computer Science, 2024, 242: 1235-1242.
5. Kim O, Park J, Park C, et al. Data asset valuation model review[J]. The Journal of Bigdata, 2021, 6(1): 153-160.
6. Deng Ganghong, Liu Le. The Limitations of Data Asset Market Pricing Mechanism and Its Rule of Law Concept[J]. Journal of Shanghai University (Social Science Edition), 2024, 41(4): 1-13.