SHAP_SVD: Integrating SHAP Values with Singular Value Decomposition for Conceptual Analysis in Regression

Yukari Shirota
Faculty of Economics
Gakushuin University
Tokyo, Japan
e-mail: yukari.shirota@gakushuin.ac.jp

Tamaki Sakura
Nikkei Economics Centre

Tokyo, Japan
e-mail: sakura@jcer.or.jp

Abstract— **In this paper, we propose a novel explainable AI (XAI) method for regression analysis, named SHAP_SVD, which integrates SHAP (SHapley Additive exPlanations) values with Singular Value Decomposition (SVD) to uncover latent structures in model interpretations. The Shapley value, a concept initially introduced by Lloyd Shapley in the field of cooperative game theory, has recently gained substantial attention in the AI community, particularly through its adaptation into the SHAP framework by Scott Lundberg. SHAP values enable us to interpret the contribution of each explanatory variable to a specific prediction by treating the prediction process as a game in which variables are players, and their marginal contributions are evaluated across all possible coalitions. In regression analysis, SHAP values can be seen as a matrix of attributions: for each observation, the contribution of each feature is calculated relative to a baseline. Our SHAP_SVD method applies SVD to this SHAP value matrix, thereby reducing dimensionality while preserving key information. The eigenvalues and corresponding eigenvectors obtained from SVD allow us to identify "concepts" or "latent semantic structures" that govern the interaction between features and the target variable. These concepts are encoded in both the left singular vectors. As a case study, we conducted a regression analysis of stock price growth rates for leading Indian and Japanese automobile manufacturers. The SHAP values were computed using a tree-based ensemble regression model, and our SHAP_SVD method was applied to reveal underlying structures. Two principal components emerged from the decomposition. In the extended analysis, we present a more detailed examination of the SHAP value distribution and structure. Specifically, we analyze how the SHAP captures nonlinear interactions that are invisible in traditional raw data correlation matrices. By comparing the performance of models evaluated using raw input features with those interpreted through SHAP, we demonstrate that SHAP-based interpretation yields greater stability, clarity, and interpretability, particularly in cases where multicollinearity or redundant variables obscure the true contribution of each feature. Our results demonstrate that SHAP analysis not only enhances the transparency of complex models but also, when combined with SVD, offers a powerful tool for discovering and visualizing conceptual dimensions underlying the data. The SHAP_SVD approach thus serves both as a diagnostic tool for regression models and a framework for semantic exploration in high-dimensional datasets.**

Keywords- Explainable AI (XAI); SHAP values; Shapley theory; Singular Value Decomposition (SVD); Dimensionality reduction; Regression analysis; Latent semantic structure; Stock price analysis; Feature attributions.

I. INTRODUCTION

In this paper, we introduce a new Explainable AI (XAI) method named SHAP_SVD, which was first introduced in [1]. XAI has emerged as a critical field, bridging the gap between complex machine learning models and human interpretability. Among the numerous XAI techniques developed, Shapley values, introduced by Lloyd Shapley, have gained prominence for their ability to allocate the contribution of each feature in a model's predictions [2-4]. Adapted into the SHAP (SHapley Additive exPlanations) framework by Lundberg [5-7], this method has become a widely used tool for interpreting machine learning models, particularly in regression analysis [8-11]. In regression analysis, SHAP values quantify the contribution of each explanatory variable to the target value by utilizing characteristic functions of the data. These values offer deep insights into feature importance and interaction. However, as the complexity and dimensionality of the data increase, interpreting SHAP values becomes increasingly challenging. Traditional SHAP methods are limited in their ability to reveal underlying structures within the data, especially when dealing with high-dimensional or multi-faceted variables.

To address this limitation, we propose a novel method, SHAP_SVD, which applies Singular Value Decomposition (SVD) to the SHAP value matrix [1]. SVD, a well-known dimensionality reduction technique, captures the core structure of a matrix by decomposing it into eigenvalues and eigenvectors [12-14]. This allows us to extract latent semantic concepts or "principal components" from the SHAP matrix. By leveraging SVD, three matrices can be obtained, which are U, Σ, and V^T (see Figure 1). This SHAP_SVD uncovers these underlying concepts, represented by two sets of eigenvectors—one from the UΣ matrix and the other from the ΣV^T matrix—thus providing a richer understanding of the relationships between explanatory variables and the target variable.

As a concrete example, we apply SHAP_SVD to the regression analysis of stock price growth rates for Indian and

Japanese automakers, using the market capitalization growth rates as the target variable. Through this analysis, we identified two key latent concepts, which we refer to as (1) Balanced (Well-balanced) type, and (2) Sales Growth Rate (SGR)-driven type," extracted from the SHAP_SVD decomposition. By plotting company data on a two-dimensional plane defined by these two principal component axes, we can conduct a detailed analysis of the characteristics driving market capitalization growth for each company. This approach enables us to visualize and understand the underlying factors that influence the stock price performance of companies in both markets.

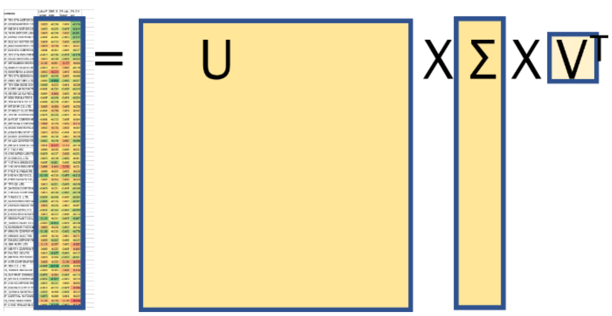

Figure 1. SVD of the given matrix.

The remainder of this paper is organized as follows. In Section II, we describe the data used for the analysis, including the sources of the data. Section III explains the methods applied, introducing both the SHAP analysis and our proposed SHAP_SVD method. Section IV presents the SHAP results, analyzing the contributions of explanatory variables to the target values. Section V details the SHAP_SVD method, illustrating how Singular Value Decomposition is applied to the SHAP matrix and how latent concepts are extracted. In Section VI, we discuss existing work related to explainable AI and dimensionality reduction, comparing these approaches with our proposed method. Section VII provides a discussion of the results and their implications. Finally, Section VIII concludes the paper with a summary of contributions and suggestions for future research.

II. DATA

In this section, we shall explain the regression data. In the regression, we use Market Capitalization (MC) data. MC amount is a stock price times the number of issued stocks. The target variable is the Indian and Japanese automakers' "**annual MC growth rates**" in 2022. The MC growth rate in year XXX is defined as ***(MC_XXX – MC_(XXX-1)) / (MC_(XXX-1))***, namely, the ratio based on the previous year. We would like to identify the dominant factors contributing to the rapid growth rates of MC. The MC data used were retrieved from the ORBIS company database by Bureau van Dijk, with the last data update date being June 22, 2024. Measure the Indian automakers' MC changes. The damages caused by COVID-19 have exposed vulnerable supply chains in the automotive industry. This regression framework assumes that the competence of supply chains and new market development are prerequisites for the long-term sustainability of companies' high business performance, leading to high stock price evaluation [15]. Therefore, we select four managerial factors as the explanatory variables. Sales Growth Ratio (**SGR) represents the new market development competence**, and **FArate** represents the supply chain competence [15-17]. The Tangible Fixed Asset amount (FA) is the third explanatory variable used to identify the impact of the firm's scalability. These factors allow companies to earn satisfactory levels of profitability, such as their stock prices, ROE, and ROA. In addition, we focus on labor productivity. Labor productivity in the manufacturing sector refers to the goods or value one worker produces within a specific period. It is a crucial metric for assessing the efficiency and competitiveness of a manufacturing operation. We aim to evaluate which factor is more significant for the target tangible assets. Labor productivity was calculated here using the following formula:

$$\begin{aligned} \textit{Labor Productivity} &= \frac{\textit{Total Value Added}}{\textit{Number of Workers}} \\ &= \frac{\textit{Net Sales} - \textit{Cost of Goods Sold}}{\textit{Number of Workers}} \end{aligned}$$

The managerial index data of the automobile companies were also retrieved from the ORBIS company database by Bureau van Dijk, with the last data update date being 2024/06/22. After removing companies with missing annual data, the number reached 67, including 11 Indian and 56 Japanese automakers. We conducted the regressions with the data.

III. METHODS

In this section, the methods we used are described. The flow chart of the analysis is as follows:

1. **XGBoost Regression**: The given data is input into the XGBoost Regressor [18], and then the regression function f(X) is generated as output.
2. **SHAP Evaluation**: Based on the regression function f(X), SHAP values for each data are calculated. In this case study, we use four explanatory variables and 67 companies, resulting in a SHAP matrix of size 67 x 4.
3. **SVD of SHAP Matrix**: Applying Singular Value Decomposition (SVD) to the SHAP matrix M, the decomposition outputs three matrices such that $M=U\Sigma V^T$
4. **SHAP_SVD Interpretation**: The eigenvectors and eigenvalues extracted from SVD are interpreted to uncover underlying concepts. The two sets of eigenvectors are referred to as CompanyEigenVectors and SHAP_Eigenvectors, representing two different viewpoints of the underlying concepts.

SHAP (SHapley Additive exPlanations) is a method based on Shapley values from cooperative game theory, designed to explain machine learning model predictions, including those in regression tasks. A key strength of SHAP is its ability to create a characteristic function for the data, allowing it to calculate the contribution of each feature based on the characteristics of individual data points. This ensures that the contribution of each feature to the model's output is computed fairly and additively. SHAP enhances the interpretability of complex models, offering insights into how specific data characteristics influence predictions. We used XGBoost as the regression algorithm.

IV. SHAP RESULTS

In this section, the results of the SHAP evaluation are presented.

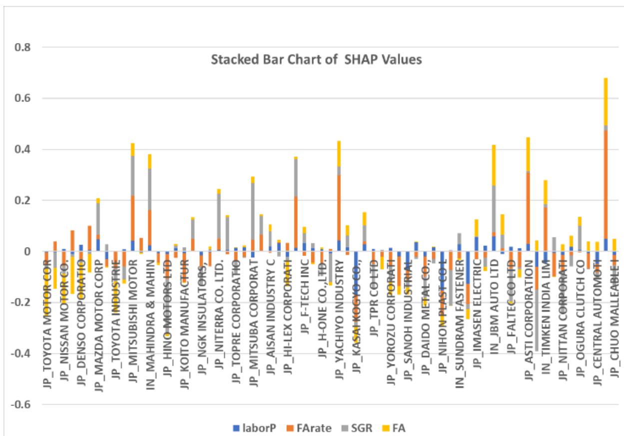

Figure 2. Stacked bar chart of the SHAP values.

The regression model, developed using XGBOOST, achieved an R-squared value exceeding 0.99, indicating a highly accurate fit to the data. Using the regression model f(X), the characteristic function is approximately evaluated. SHAP values are found based on the characteristic function. Figure 2 shows the SHAP values. The horizontal line shows the company IDs. Figure 2 illustrates a stacked bar chart of SHAP values of the individual companies. Each company has four SHAP values corresponding to the four explanatory variables.

The horizontal zero line in Figure 2 shows the average target value of the companies. Companies with positive values tend to be above-average performers. On the other hand, companies with negative values are evaluated as lower-performing than average. In corporate analysis, a fundamental question is whether a company performs above or below the industry standard. This is often the first benchmark used in evaluating a firm's standing. Therefore, the SHAP framework, which sets the industry average contribution to zero, aligns well with this evaluation logic. From a management science perspective, this interprets SHAP values as highly intuitive and meaningful.

The sum of four SHAP values in each company becomes its deviation from the target value. Table 1 shows the correlation coefficients between the target and the individual SHAP values. The highest correlation is between the FA rate_SHAP and the target, with a correlation coefficient of 0.84. The second highest is between SGR_SHAP and the target, with a value of 0.73. The third highest is between FA_SHAP and the target, with a value of 0.70. The result says that the most dominant factor for the target is FArate. Then, the second one is SGR.

Table 1. Correlation coefficients between the target and SHAP values.

	target	laborP_SHAP	FArate_SHAP	SGR_SHAP	FA_SHAP
target	1.00				
laborP_SHAP	0.55	1.00			
FArate_SHAP	0.84	0.36	1.00		
SGR_SHAP	0.73	0.24	0.41	1.00	
FA_SHAP	0.70	0.30	0.48	0.27	1.00

Figure 3 shows the relationship between FA_SHAP and the target as a scattering plot, and Figure 4 shows that between SGR_SHAP and the target. The SHAP correlation becomes higher than that between a raw feature value and the target.

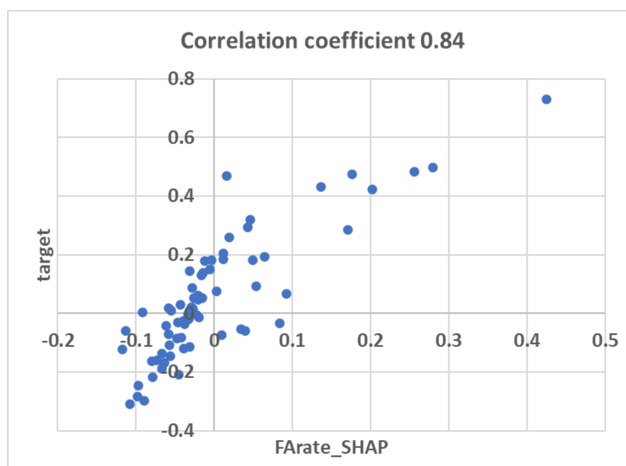

Figure 3. Scattering plot between FArate_SHAP and the target.

One key advantage of SHAP values over raw feature values is their ability to capture the actual influence of features on the target, even when simple correlations are weak. While a raw feature may show low correlation with the target variable, its corresponding SHAP value—reflecting its marginal contribution within the model—can exhibit a much stronger association. This occurs because SHAP values incorporate interactions and conditional dependencies that the model has learned. As a result, SHAP-based correlation analysis provides a more accurate representation of how features influence model predictions, particularly in complex, nonlinear, or high-dimensional settings.

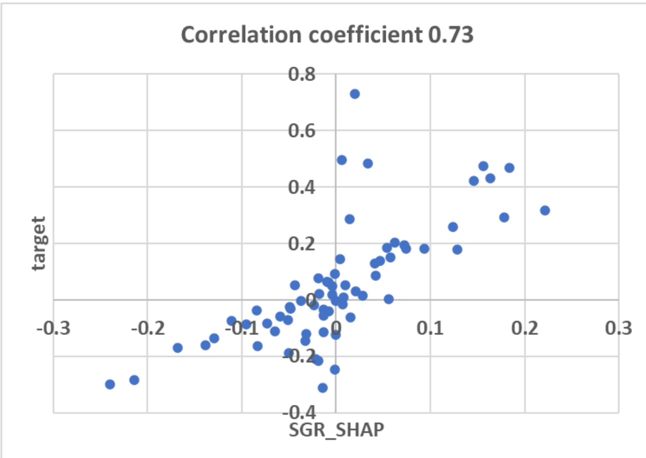

Figure 4. Scattering plot between SGR_SHAP and the target.

Another representation of the same set of SHAP values as in Figure 2 is the SHAP matrix M shown in Table 2. The matrix size becomes 67 times 4. The line represents a company, and each row displays an SHAP value. The matrix is divided by SVD.

Table 2. SHAP value matrix.

company	laborP_SHAP	FArate_SHAP	SGR_SHAP	FA_SHAP
JP_TOYOTA MOTO	-0.003	-0.045	-0.021	-0.188
JP_HONDA MOTOR	-0.004	0.039	-0.060	-0.083
JP_NISSAN MOTOR	0.010	-0.056	-0.033	-0.116
IN_TATA MOTORS	-0.009	0.084	-0.013	-0.144
JP_DENSO CORPO	0.026	-0.031	-0.013	-0.146
JP_SUZUKI MOTOR	0.007	0.092	-0.010	-0.073
JP_MAZDA MOTOR	0.047	0.019	0.124	0.019
JP_SUBARU CORP	-0.031	-0.027	0.028	-0.005
JP_TOYOTA INDUS	0.000	-0.067	-0.050	-0.122
JP_ISUZU MOTORS	0.008	-0.059	-0.051	-0.016
JP_MITSUBISHI MC	0.041	0.176	0.156	0.050
IN_MARUTI SUZUK	-0.004	0.054	-0.001	-0.006
IN_MAHINDRA & M	0.024	0.137	0.163	0.057
JP_TOYOTA BOSHO	-0.008	-0.034	0.000	-0.009
JP_HINO MOTORS	-0.009	-0.098	-0.214	-0.010
JP_TOYODA GOSEI	0.012	-0.026	0.010	0.007
JP_KOITO MANUFA	-0.007	-0.113	0.015	-0.004
IN_ASHOK LEYLAN	-0.001	0.049	0.074	0.011
JP_NGK INSULATO	-0.016	-0.043	-0.073	0.001
JP_TOKAI RIKA CO	-0.004	-0.055	0.007	0.013
JP_NITERRA CO. LT	0.005	0.043	0.179	0.016

V. SHAP_SVD METHOD

In this section, SHAP_SVD method is explained. After the SVD of the SHAP matrix, the singular value (SV) lists are obtained as the diagonal elements in matrix Σ (see Figure 1).

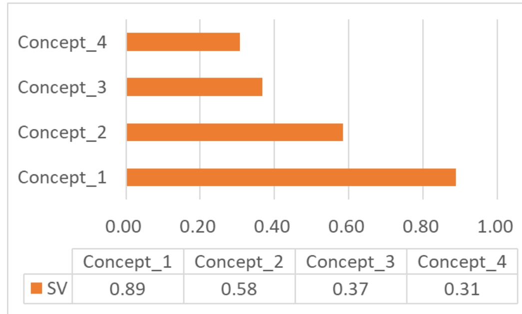

Figure 5. Singular values of the SHAP matrix M.

The SVs express, as shown in Figure 5, the strength of the latent concepts. The ratio is approximately 9:6:4:3. The eigenvalue of each concept becomes the square of the SV mathematically. Then the eigenvalues are those shown in Figure 6. The percentages of eigenvalues are shown in Table 3. Approximately 58% of the concepts can be explained by concept 1. The second concept has about 25%. Therefore, an overview of the structure can be captured simply by examining the first and second concepts.

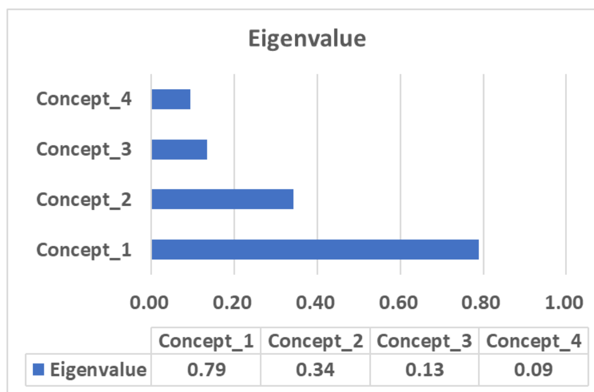

Figure 6. Eigenvalues of the matrix $U\Sigma$.

Table 3. Eigenvalue percentage of the concepts.

Concept_1	Concept_2	Concept_3	Concept_4
0.79	0.34	0.13	0.09
58%	25%	10%	7%

Then, we will interpret the meaning of the concepts, focusing on the two largest SVs. Two matrices are pre-constructed through matrix multiplication in advance: $U\Sigma$ and ΣV^T (see Figure 7).

The individual concept can be represented by two expressions: CompanyEigenVectors and SHAP_EigenVectors as shown in Figure 7. The terminology used here was defined specifically for this case study and does not represent a general or standardized name. In other contexts, it could be referred to as *DateEigenVector* or other case-appropriate labels.

For example, concept 1 is defined by the first row of $U\Sigma$, and is represented by the first line of ΣV^T (see Figure 7).

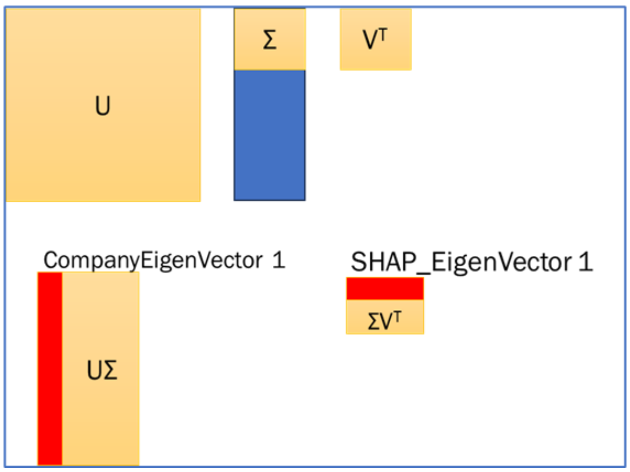

Figure 7. CompanyEigenVectors and SHAP_EigenVectors in SVD.

First, using SHAP_EigenVectors, the concepts will be expressed in Figure 8. The bottom bar graph presents SHAP_EigenVector_1, which has four elements. The upper graph shows SHAP_EigenVector_2, which has four elements as well.

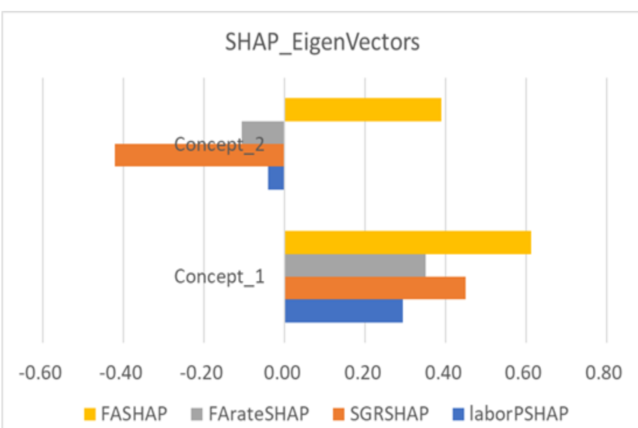

Figure 8. SHAP elements of Eigenvalues 1 and 2.

Our interpretation of the two concepts is as follows:

- 1st concept: All elements cooperate and have high values, with particularly high SHAP for FA.
- 2nd concept: SGR_SHAP is high, and FA's SHAP is low (expressing it this way reverses the sign of the vector elements).

As a result, we name the two concepts as
(1) Balanced (Well-balanced) type and
(2) Sales Growth Rate (SGR)-driven type.

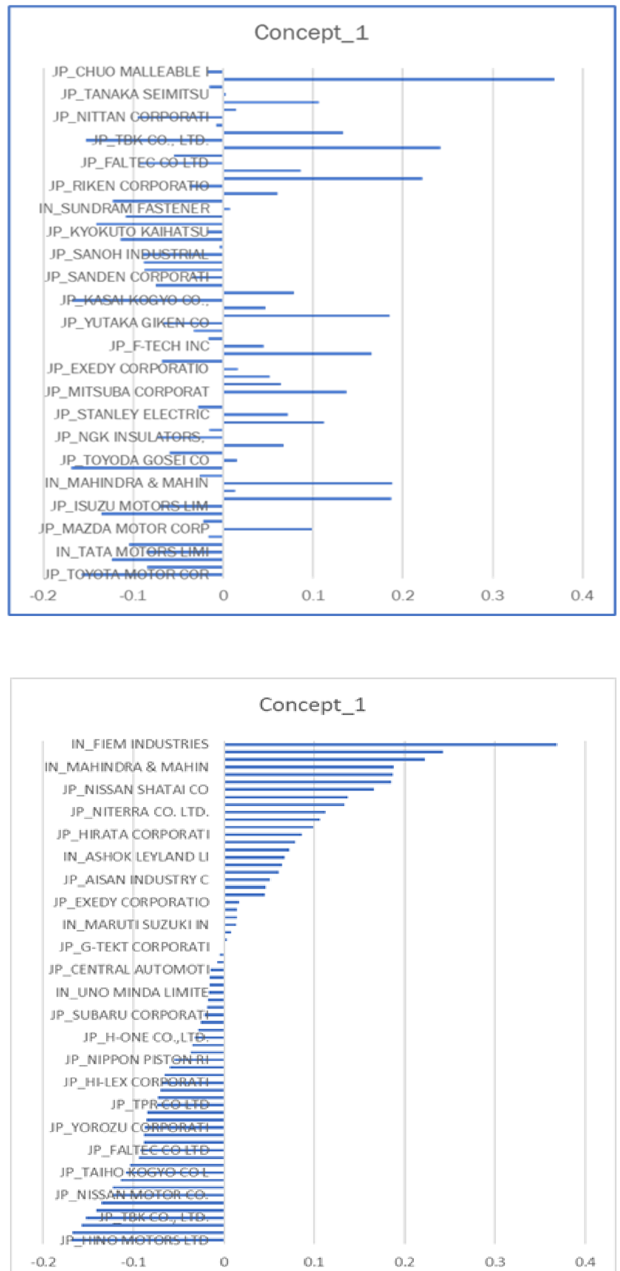

Figure 9. CompanyEigenVector_1 (the bottom is the sorted one).

Then, using CompanyEigenVectors, the concepts will be interpreted. Figure 9 shows CompanyEigenVector_1. The bottom graph is the sorted version. The largest element company was FIEM Industries. FIEM is a well-established company in India, primarily known for its expertise in automotive lighting. With over 50 years of experience, FIEM has grown into a leading supplier for original equipment manufacturers (OEMs) in India and abroad. In the representation using CompanyEigenVectors, the first concept can be interpreted as companies with SHAP distributions similar to FIEM.

Figure 10 shows CompanyEigenVector_2. The lowest value company is JP_NISSAN SHATAI, and the third lowest company is TOYOTA. The highest value company is JP_CHUO MALLEABLE.

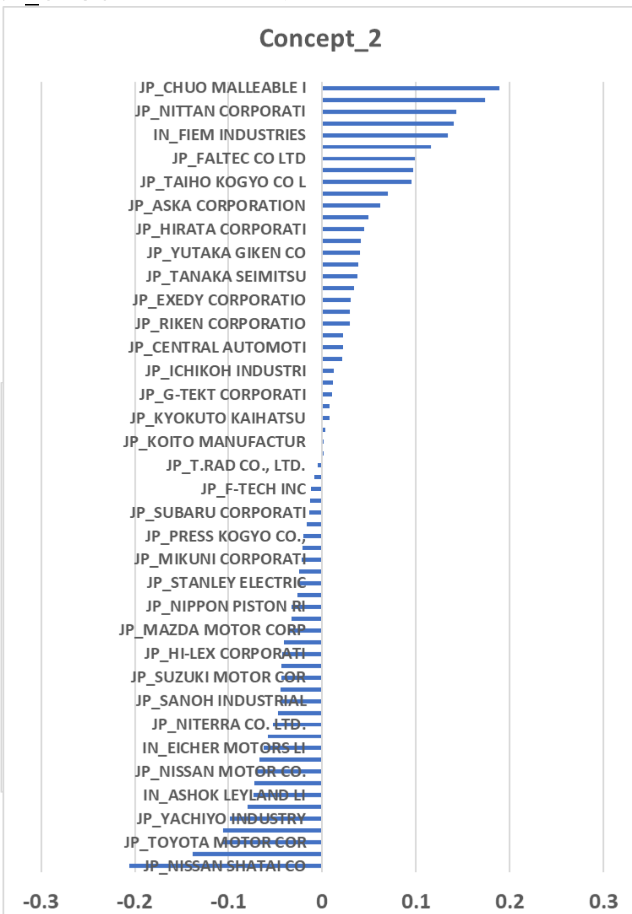

Figure 10. CompanyEigenVector_2.

Figure 11 shows a scattering plot of CompanyEigenVector_1 elements and CompanyEigenVector_2 element values of 67 companies. The x-axis represents the concept 1 level, and the y-axis represents the concept 2 level.

First, the five companies with the highest CompanyEigenVector_1 values will be explained in Figure 12, along with their corresponding SHAP values. The FIEM and ASTI are included there. The five companies are well-balanced, and the target deviation values are positive, with higher values. Seeing these five companies' SHAP distributions, the ratio of FA_SHAP and SGR_SHAP are higher than others, which are the same as the concept 1 representation in Figure 8; SHAP_EigenVector 1.

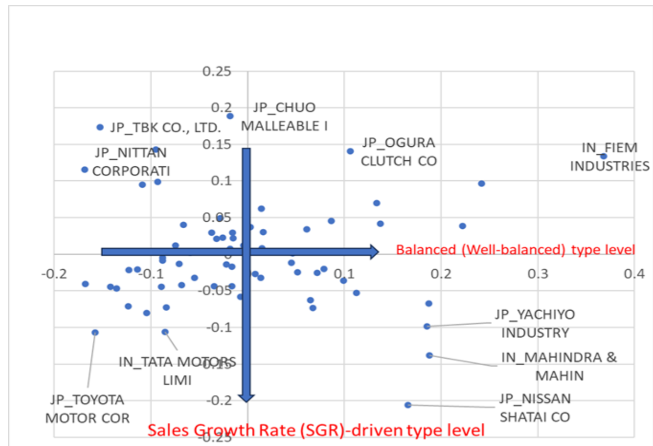

Figure 11. A scattering plot of CompanyEigenVector_1 elements and CompanyEigenVector_2 element values of 67 companies.

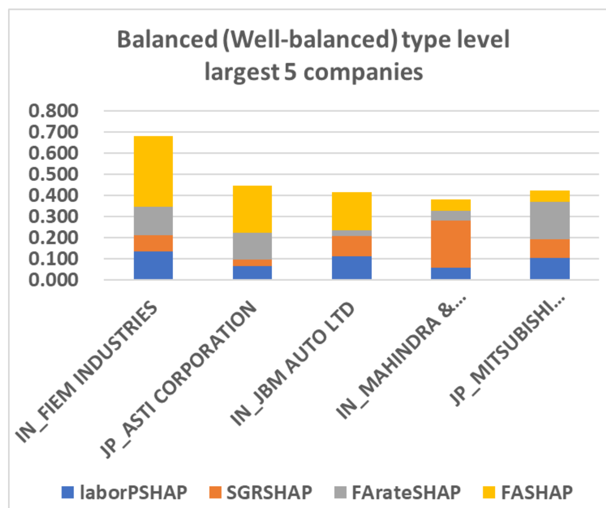

Figure 12. SHAP values of the Balanced type level of the five largest companies.

On the other hand, the five companies with the lowest CompanyEigenVector_1 in Figure 13 have negative target deviation values. The distribution of SHAP values on each company represent the company characteristics. From the comparison between Figures 12 and 13, a company with higher concept 1 value tends to be higher in the target value.

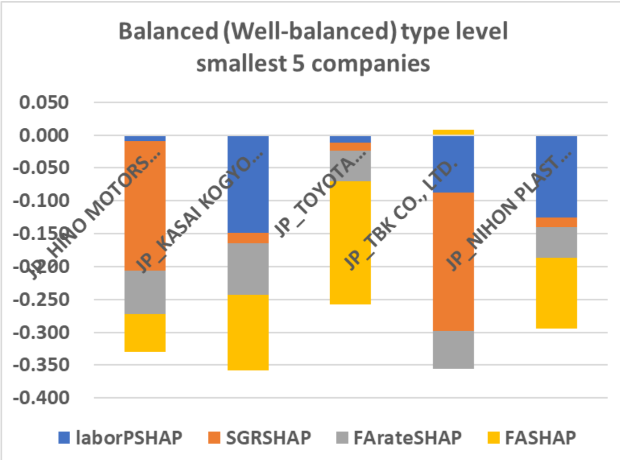

Figure 13. SHAP values of the Balanced type level of the five lowest companies.

Next, we will see SGR-driven type concept 2. Figure 14 illustrates the SHAP values of the top five companies in the SGR-driven type level. As shown in Figure 11, the second concept (y-axis) is oriented downwards, and the company with the highest y-value is Nissan Shatai, followed by Mahindra, Toyota, Tata, and Yachiyo. Figure 14 shows these five companies' stacked bar chart of SHAP values.

The target deviation positive three companies are SGR-driven ones. On the other hand, the other two companies are anti-SGR driven and have lower FA_SHAP values. As shown in Figure 8, concept 2 exhibits a higher SGR_SHAP and lower FA_SHAP. There is a reverse relationship between SGR and FA. Therefore, these anti-SGR-driven companies show the feature of concept 2, and these are evaluated as higher value companies concerning concept 2.

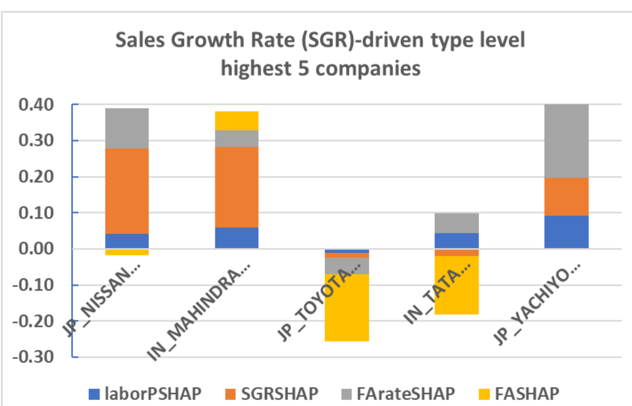

Figure 14. SHAP values of the SDG-driven type level of the five largest companies.

Next, we will evaluate the SHAP values of the five smallest companies of the SDG-driven type level (see Figure 15). In the three companies with positive SHAP values, FA_SHAP (Fixed Assets SHAP) is large, while SGR_SHAP is small, indicating that these companies are driven more by the size of their tangible assets rather than sales growth. For the two companies with negative target deviation values, SGR_SHAP is dragging performance less than FA_SHAP, with FA_SHAP values being close to zero. This suggests that FA has minimal impact on these two companies.

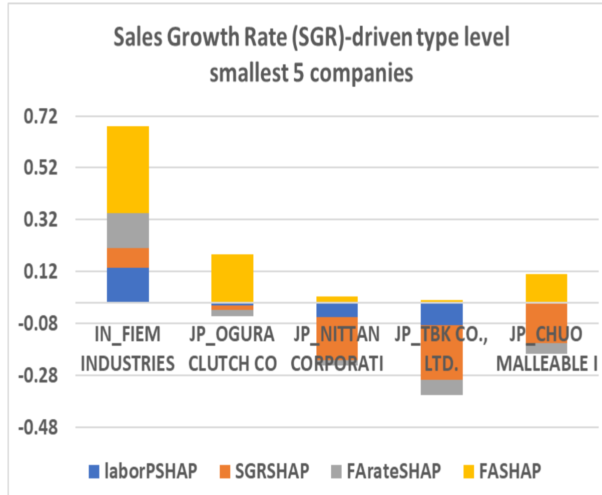

Figure 15. SHAP values of the SDG-driven type level of the smallest five companies.

To summarize the findings thus far, **concept 1 shows a positive relationship with the target deviation**. Companies with higher values of concept 1 also tend to exhibit higher target deviations. In contrast, **concept 2 exhibits mixed signs in its component values**, depending on the relative dominance of different SHAP contributions. This variation arises from the structure of concept 2 itself, which is characterized by a **positive loading on SGR_SHAP** and a **negative loading on FA_SHAP**. As a result, the overall direction of concept 2 depends on which factor exerts a stronger influence in a given company's SHAP profile.

VI. Existing Work

In this section, the related existing works are presented. The first allocation field involves stock price evaluation, and the second allocation field utilizes text mining.

(1) Random Matrix Theory (RMT) and portfolio

RMT has been applied to stock market analysis to reduce noise in financial data. RMT helps distinguish real market signals from random fluctuations in stock price correlations [19-24]. The flow charts of the method are as follows:

1. Correlation Matrix: Begin by calculating the correlation matrix of stock returns. This matrix sizes the number of companies times the number of sales dates.
2. RMT Filtering: RMT is used to separate meaningful signals from random noise. Eigenvalues of the correlation matrix are compared with theoretical RMT

predictions. Larger eigenvalues represent accurate market information, while smaller ones reflect noise.

3. SVD (Singular Value Decomposition): SVD is applied to clean the correlation matrix further, focusing on the significant components. This improves the matrix's accuracy, filtering out noise.
4. Portfolio Optimization: Using the noise-reduced correlation matrix, more accurate risk and return estimates can be made, improving portfolio construction.

(2) Latent Semantic Analysis (LSA)

LSA is a widely used technique in natural language processing (NLP), primarily for analyzing semantic relationships between documents. It is often applied in tasks such as topic modeling, semantic analysis, and information retrieval [25-31].

Overview of LSA:

1. Purpose: LSA aims to convert the semantic relationships between words and documents into a lower-dimensional latent semantic space, allowing for the identification of similarities and relationships between documents. This helps uncover hidden patterns or topics within the text.
2. Method: LSA begins by creating a co-occurrence matrix that captures how often words appear together in a document. This matrix models the relationships between words and documents. Then, SVD is applied to reduce the dimensionality of the matrix. By using SVD, LSA compresses the high-dimensional data while preserving the important semantic relationships and filtering out noise.

LSA is a powerful mathematical approach for interpreting the semantic structure of text and is utilized in search engines, automatic summarization systems, document clustering, and more. SVD techniques are mathematically explained in [12, 14]. The two kinds of eigen vectors and the relationship among the three decomposed matrices are clearly explained using visualization in [32].

VII. Discussion

In this section, we will discuss the result. The objective of the analysis is to group companies. The proposed SHAP_SVD method can extract the essence of the given SHAP value matrix.

In the case of India and Japan automakers, the two extracted concepts were obtained: (1) Balanced (Well-balanced) type, and (2) Sales Growth Rate (SGR)-driven type. The concept can be represented as two aspects in this case that are CompanyEigenVectors and SHAP_EigenVectors. There are found as the results of SVD.

Using each CompanyEigenVectors' element values, we can measure each company's (1) Balanced (Well-balanced) type level and (2) Sales Growth Rate (SGR)-driven type level. As shown in Figure 11, the scattering plot of the companies, generated by CompanyEigenVectors, can reveal the individual characteristics of each company. The horizontal axis represents the Balanced (Well-balanced) type level. These higher-level companies can be divided along the vertical axis into two groups: an "SGR_SHAP higher and FA_SHAP lower" group and a "FA_SHAP higher and SGR_SHAP lower" group. This means that these companies exhibit a similar pattern of feature contributions, reflecting a particular type of balance or focus in their business models.

SHAP values can more accurately reflect each company's characteristics than using the raw input data. Therefore, analyzing SHAP values through SVD (Singular Value Decomposition) allows for more accurate dimensionality reduction based on the characteristics of each company. This method enhances the ability to capture distinct business drivers by compressing the data in a way that aligns with each company's unique attributes, offering deeper insights compared to standard SHAP analysis.

In corporate management, creating appropriate KPIs (Key Performance Indicators) is crucial. The EigenVectors (principal component axes) derived from SHAP_SVD analysis can serve as the first step in developing these KPIs. By identifying the most critical factors influencing business performance through dimensionality reduction, SHAP_SVD helps to highlight key metrics that align with a company's unique characteristics, providing a strong foundation for effective KPI creation.

Although the present case study focuses on four managerial indicators, the SHAP_SVD framework is model-agnostic and can be directly applied to high-dimensional settings. Because SHAP_SVD operates on the SHAP attribution matrix rather than the raw input space, it can handle models that generate large attribution vectors, including those used for image, audio, or text data. In contrast, traditional neural-fuzzy approaches rely on human-interpretable rule bases, and the number of rules increases exponentially as the number of input variables grows, making high-dimensional applications structurally difficult. Therefore, SHAP_SVD provides a more flexible and scalable approach for extracting latent concepts from modern machine-learning models.

VIII. Conclusions

In this paper, we propose a novel explainable AI (XAI) methodology named SHAP_SVD, which integrates SHAP values with Singular Value Decomposition (SVD) to extract latent semantic structures from regression models. This approach addresses a key challenge in modern machine learning applications—the difficulty of interpreting complex models in high-dimensional settings.

While SHAP values already offer a theoretically sound and practically beneficial way to understand feature contributions, they often lack the structural abstraction needed for global interpretation and cross-sample comparison. Our SHAP_SVD method fills this gap by applying dimensionality reduction to the SHAP value matrix, revealing underlying concepts that govern the feature-target relationships.

In our case study on stock price growth rates of Indian and Japanese automakers, the SHAP_SVD analysis revealed two principal latent concepts:

(1) a Balanced (Well-balanced) type, characterized by uniform positive contributions across features, especially Fixed Assets (FA), and

(2) an SGR-driven type, emphasizing Sales Growth Rate (SGR) over other explanatory variables.

These concepts were visualized using the eigenvectors derived from the SHAP matrix decomposition, providing two complementary perspectives CompanyEigenVectors, which characterize companies, and SHAP_EigenVectors, which describe feature patterns. This framework offers several significant contributions: First, SHAP_SVD enhances interpretability by translating local explanations (SHAP values per sample) into global structures (principal components), allowing researchers and analysts to detect patterns and clusters among entities. Second, it provides a foundation for data-driven KPI (Key Performance Indicator) design, where companies can be evaluated along meaningful conceptual axes extracted from model behavior, not just raw data.

Moreover, SHAP_SVD's utility is not limited to regression problems in the automotive sector and company sector. Its generality suggests broad applicability in other domains such as finance, healthcare, education, and policy evaluation—any field in which explainable AI meets high-dimensional structured data. For instance, medical diagnostics models could benefit from latent factor discovery among symptoms and test results, or macroeconomic models could use SHAP_SVD to explain and classify country-level performance indicators.

Future work will explore the following directions:

- Extending SHAP_SVD to classification tasks, where class probability contributions rather than regression outputs are analyzed.
- Combining SHAP_SVD with clustering algorithms to automatically group entities based on shared conceptual characteristics.
- Developing interactive visualization tools, enabling stakeholders to navigate the concept space in a user-friendly manner.

In conclusion, SHAP_SVD is a promising tool for conceptual interpretation of machine learning models. By bridging SHAP's local attribution with SVD's global structure extraction, it provides a robust and interpretable framework for analyzing the inner workings of predictive models. We believe this methodology offers not only academic value but also practical utility for decision-makers aiming to derive actionable insights from complex datasets.

Acknowledgment

The research was partly supported by the fund of Gakushuin University Computer Centre project 2024 and Gakushuin University Economics and Management project 2025.

References

[1] Y. Shirota and T. Sakura, "Exploring Latent Concepts in SHAP Values -A New Approach Using Singular Value Decomposition -," in *DBKDA 2025*, Lisbon, Portugal, 2023/03/9-13 2025: IARIA XPS Press, pp. 1-6, doi: 978-1-68558-244-9. [Online]. Available: https://www.thinkmind.org/library/DBKDA/DBKDA_2025/dbkda_2025_1_10_58001.html (accessed 2025/11/14).

[2] A. E. Roth, "Introduction to the Shapley Value," *The Shapley value,* pp. 1-27, 1988.

[3] A. E. Roth, *The Shapley Value: essays in honor of Lloyd S. Shapley*. Cambridge University Press, 1988.

[4] L. S. Shapley, "A value for n-Person Games, Contributions to the Theory of Games, 2, 307–317," ed: Princeton University Press, Princeton, NJ, USA, 1953.

[5] S. M. Lundberg, G. G. Erion, and S.-I. Lee, "Consistent Individualized Feature Attribution for Tree Ensembles," *arXiv preprint arXiv:1802.03888,* 2018.

[6] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model predictions," *Advances in neural information processing systems,* vol. 30, 2017.

[7] S. M. Lundberg and S.-I. Lee, "Consistent Feature Attribution for Tree Ensembles," *arXiv preprint arXiv:1706.06060,* 2017.

[8] A. R. Javed, W. Ahmed, S. Pandya, P. K. R. Maddikunta, M. Alazab, and T. R. Gadekallu, "A Survey of Explainable Artificial Intelligence for Smart Cities," *Electronics,* vol. 12, no. 4, p. 1020, 2023.

[9] R. Dwivedi *et al.*, "Explainable AI (XAI): Core Ideas, Techniques, and Solutions," *ACM Computing Surveys,* vol. 55, no. 9, pp. 1-33, 2023.

[10] A. Chaddad, J. Peng, J. Xu, and A. Bouridane, "Survey of Explainable AI Techniques in Healthcare," *Sensors,* vol. 23, no. 2, p. 634, 2023.

[11] Y. Shirota, K. Kuno, and H. Yoshiura, "Time Series Analysis of Shap Values by Automobile Manufacturers Recovery Rates," in *Proceedings of the 2022 6th International Conference on Deep Learning Technologies*, 2022, pp. 135-141.

[12] C. M. Bishop and N. M. Nasrabadi, *Pattern Recognition and Machine Learning* (no. 4). Springer, 2006.

[13] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2009.

[14] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*. Elsevier, 2006.

[15] Y. Shirota, M. Fujimaki, E. Tsujiura, M. Morita, and J. A. D. Machuca, "A SHAP Value-Based Approach to Stock Price Evaluation of Manufacturing Companies," in *2021 4th International Conference on Artificial Intelligence for Industries (AI4I)*, 2021: IEEE, pp. 75-78.

[16] M. Fujimaki, E. Tsujiura, and Y. Shirota, "Automobile Manufacturers Stock Price Recovery Analysis at COVID-19 Outbreak," in *6th World Conference on Production and Operations Management – P&OM Nara 2022*, Nara, Japan, 2022: EurOMA (European Operations Management Association), pp.1-6, Decision Science Institute Best Paper Award.

[17] K. Yamaguchi, "Relationship Analysis Between Stock Prices and Financial Statements in the Automobile Industry," in *2023 14th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)*, 2023: IEEE, pp. 442-445.

[18] XGBoostDevelopers. "XGBoost Documentation (Revision 534c940a.)." Available: https://xgboost.readthedocs.io/en/stable/ (accessed 2025/11/14).

[19] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, T. Guhr, and H. E. Stanley, "Random Matrix Approach to Cross Correlations in Financial Data," *Physical Review E,* vol. 65, no. 6, p. 066126, 2002.

[20] M. Potters, J.-P. Bouchaud, and L. Laloux, "Financial Applications of Random Matrix Theory: Old laces and new pieces," *arXiv preprint physics/0507111,* 2005.
[21] A. Utsugi, K. Ino, and M. Oshikawa, "Random Matrix Theory Analysis of Cross Correlations in Financial Markets," *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics,* vol. 70, no. 2, p. 026110, 2004.
[22] Z. Bai, H. Liu, and W. K. Wong, "Enhancement of the Applicability of Markowitz's Portfolio Optimization by Utilizing Random Matrix Theory," *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics,* vol. 19, no. 4, pp. 639-667, 2009.
[23] H. Aoyama, Y. Fujiwara, Y. Ikeda, H. Iyetomi, and W. Souma, *Econophysics and Companies: Statistical Life and Death in Complex Business Networks*. Cambridge University Press, 2010.
[24] J. Bun, J.-P. Bouchaud, and M. Potters, "Cleaning Large Correlation Matrices: Tools from Random Matrix Theory," *Physics Reports,* vol. 666, pp. 1-109, 2017.
[25] T. K. Landauer, P. W. Foltz, and D. Laham, "An Introduction to Latent Semantic Analysis," *Discourse processes,* vol. 25, no. 2-3, pp. 259-284, 1998.
[26] N. Evangelopoulos, X. Zhang, and V. R. Prybutok, "Latent Semantic Analysis: Five Methodological Recommendations," *European Journal of Information Systems,* vol. 21, no. 1, pp. 70-86, 2012.
[27] S. T. Dumais, "Latent Semantic Analysis," *Annual Review of Information Science and Technology (ARIST),* vol. 38, pp. 189-230, 2004.
[28] T. K. Landauer, D. S. McNamara, S. Dennis, and W. Kintsch, *Handbook of Latent Semantic Analysis*. Psychology Press, 2007.
[29] S. T. Dumais, G. W. Furnas, T. K. Landauer, S. Deerwester, and R. Harshman, "Using Latent Semantic Analysis to Improve Access to Textual Information," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1988, pp. 281-285.
[30] Y. Shirota and B. Chakraborty, "Visual Explanation of Eigenvalues and Math Process in Latent Semantic Analysis," *Information Engineering Express,* vol. 2, no. 1, pp. 87-96, 2016.
[31] Y. Shirota and B. Chakraborty, "Visual Explanation of Mathematics in Latent Semantic Analysis," in *2015 IIAI 4th International Congress on Advanced Applied Informatics*, 2015: IEEE, pp. 423-428.
[32] Y. Shirota and B. Chakraborty, "Visual Explanation of Eigenvalues and Math Process in Latent Semantic Analysis," *Information Engineering Express, Information Engineering Express,* vol. 2, no. 1, pp. 87-96, 2016. [Online]. Available: https://www.iaiai.org/journals/index.php/IEE/article/view/70 (accessed 2025/11/14).