Exploring Latent Concepts in SHAP Values

- A New Approach Using Singular Value Decomposition -

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

Tmaki Sakura
Nikkei Economics Centre

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

***Abstract*— In this paper, we introduce a novel explainable AI method called "SHAP_SVD" for regression analysis. The Shapley value, originally developed by Lloyd Shapley, has gained prominence as a key tool in explainable AI (XAI) through its adaptation as SHAP by Lundberg. In regression analysis, SHAP values are computed using characteristic functions of the data, representing the contribution of each explanatory variable to the target value. Our proposed SHAP_SVD method applies Singular Value Decomposition (SVD), a dimensionality reduction technique, to the SHAP value matrix. The eigenvalues and eigenvectors extracted via SVD capture the core structure of the SHAP matrix, revealing "concepts" or "latent semantic concepts." In SVD, these concepts are represented by two sets of eigenvectors. As a case study, we demonstrate the regression analysis of stock price growth rates for Indian and Japanese automakers, where two principal concepts were identified, consistently reflected across both sets of eigenvectors.**

Keywords- ***XAI; Shapley values; SHAP; Singular Value Decomposition; India automakers.***

I. Introduction

EXplainable AI (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 [1-3]. Adapted into the SHapley Additive exPlanations (SHAP) framework by Lundberg [4-6], this method has become a widely used tool for interpreting machine learning models, particularly in regression analysis [7-10]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, the interpretation of SHAP values becomes 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. SVD, a well-known dimensionality reduction technique, captures the core structure of a matrix by decomposing it into eigenvalues and eigenvectors [11-13]. This allows us to extract latent semantic concepts or "principal components" from the SHAP matrix. By leveraging SVD, SHAP_SVD uncovers these underlying concepts, represented by two sets of eigenvectors, 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 are able to 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.

The remainder of this paper is organized as follows. In Section 2, we describe the data used for the analysis, including the data sources. Section 3 explains the methods applied, introducing both the SHAP analysis and our proposed SHAP_SVD method. Section 4 presents the SHAP results, analyzing the contributions of explanatory variables to the target values. Section 5 details the SHAP_SVD method, illustrating how Singular Value Decomposition is applied to the SHAP matrix and how latent concepts are extracted. In Section 6, we discuss existing work related to explainable AI and dimensionality reduction, comparing these approaches with our proposed method. Section 7 provides a discussion of the results and their implications. Finally, Section 8 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 find the dominant factors for the rapid MC growth rates. The MC data we used were retrieved from the ORBIS company database by Bureau van Dijk, the last data update date being 2024/06/22.

The damages caused by COVID-19 have revealed vulnerable supply chains in automakers. This regression frame assumes that the competence of supply chains and new market development are prerequisites for the long-run sustenance of companies' high business performance, leading to high stock price evaluation [14]. 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 [14-16]. 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, Return On Equity (ROE), and Return On Assets (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 want to evaluate which is more significant on the target tangible assets or labor productivity. Labor productivity was calculated here using the following formula:

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

The managerial index data of the automobile companies were also retrieved from the ORBIS company database. 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 [17], 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 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 different two viewpoints of the underlying concepts.

SHAP 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.

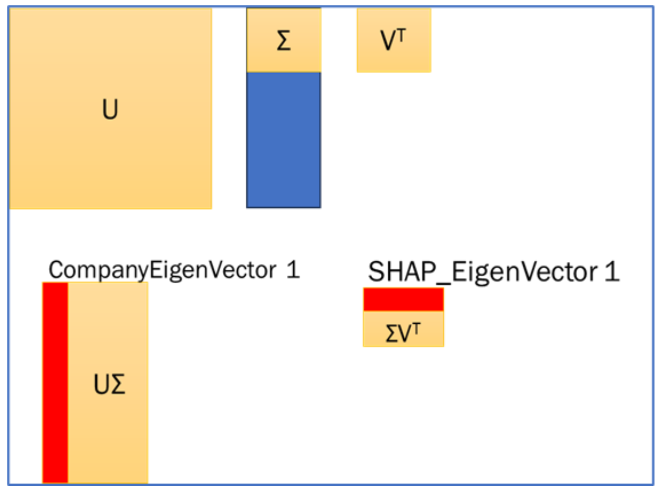

Figure 1. CompanyEigenVectors and SHAP_EigenVectors in SVD.

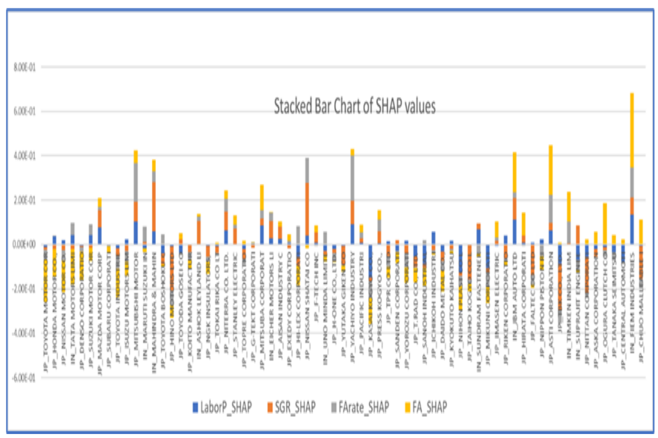

Figure 2. Stacked bar chart of the SHAP values.

In step 3, the SHAP matrix M is decomposed by SVD. As shown in Figure 1, the SHAP matrix M is decomposed to the matrixes, U, Σ, and V. The shape of U is squared.

The matrix Σ is padded by 0 in the blue area in Figure 1. The shape of U times Σ becomes a rectangle. The first row of this rectangle becomes CompanyEigenVector_1. CompanyEigenVector_1 corresponds to the first EigenVector. The shape of Σ times V^T becomes a rectangle. The first line of this rectangle becomes SHAP_EigenVector_1. SHAP_EigenVector _1 corresponds to the first EigenVector.

In step 4, SHAP_SVD interpretation, the two kinds of eigenvectors are evaluated, which are named in this case study CompanyEigenVector and SHAP_EigenVector (see the red parts in Figure 1).

IV. SHAP RESULTS

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

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 stack bar chart of SHAP values of the individual companies. There in each company has four SHAP values corresponding to the four explanatory variables. The horizontal zero line shows the average target value of the companies. The sum of four SHAP values in each company becomes its deviation of the target value from the average. The SHAP matrix M size becomes 67 times 4 which is the target matrix of the SVD.

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 (see Figure 3). The SVs express the strength of the latent concepts. The ratio is approximately 9:6:4:3.

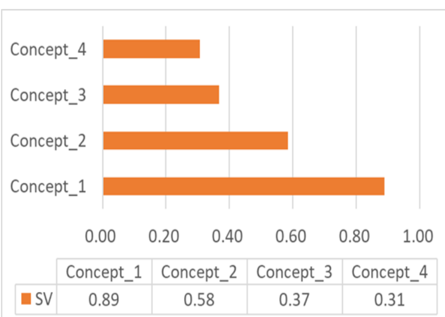

Figure 3. Singular values of the SHAP matrix M.

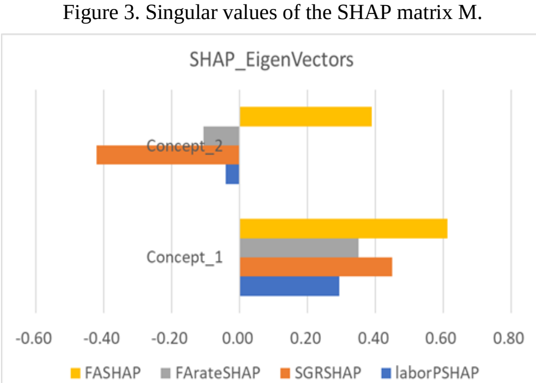

Figure 4. SHAP_Eigenvectors_1 and SHAP_Eigenvectors_2.

Then, we will interpret the meaning of the concepts, focusing on the two largest SVs. The concepts can be interpreted by the two aspects. First, using SHAP_EigenVectors, the concepts will be expressed in Figure 4. The bottom bar graph presents the SHAP_EigenVector_1 with four elements. 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).

We name the concepts (1) Balanced (Well-balanced) type and (2) Sales Growth Rate (SGR)-driven type.

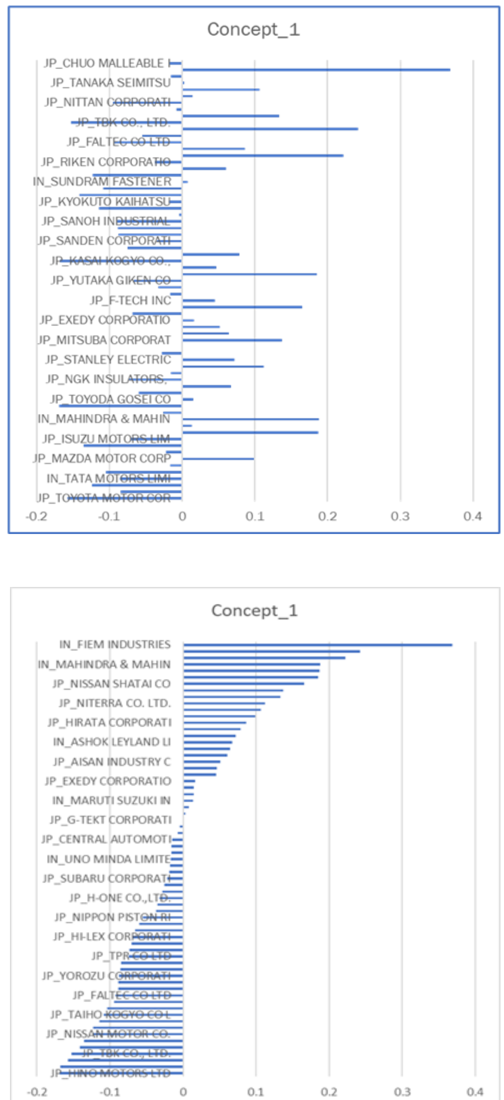

Figure 5. CompanyEigenVector_1 (the bottom is the sorted one).

Then, using CompanyEigenVectors, the concepts will be interpreted. Figure 5 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.

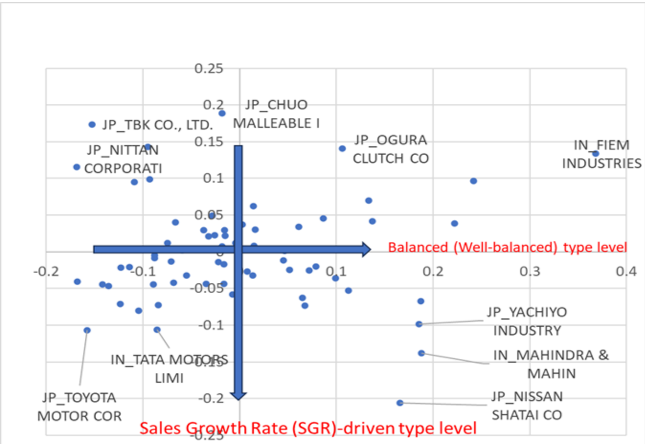

Figure 6. A scattering plot of CompanyEigenVector_1 elements and CompanyEigenVector_2 element values of 67 companies

Figure 6 shows a scattering plot of CompanyEigenVector_1 element values and CompanyEigenVector_2 element values of 67 companies. The representatives concerning CompanyEigenVector_2 elements, which measure the SGR-driven level, will be interpreted. Figure 7 shows the first SHAP values of the SGR-driven type level's highest five companies. As shown in Figure 6, the first second principle component (y-axis) is oriented downwards, and the company with the highest y-value is Nissan Shatai, followed by Mahindra as the second highest.

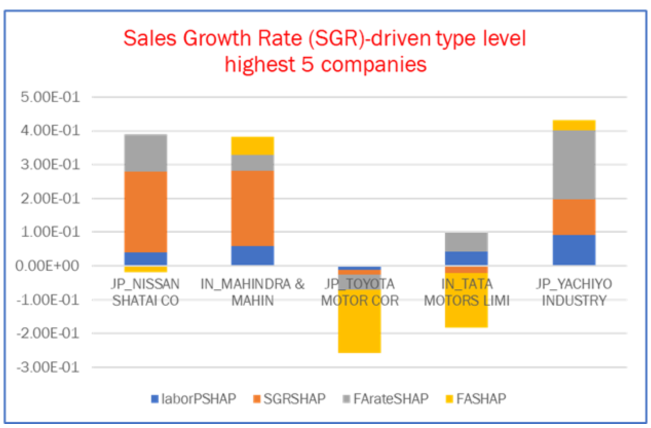

Figure 7. SHAP values of the SDG-driven type level highest five companies.

In the three companies like Nissan Shatai and Mahindra, where SHAP values are positive, the SGR_SHAP (Sales Growth Rate SHAP) is large while FA_SHAP (Fixed Assets SHAP) is small. This suggests sales growth is the main driving factor rather than the size of tangible fixed assets. For two companies with negative SHAP values, the impact of SGR_SHAP dragging performance is smaller (almost zero), compared to the negative impact of FA_SHAP.

Toyota, for example, typically the strength is tangible fixed assets, with FA_SHAP being large and positive when the target value is positive and negative when the target is negative.

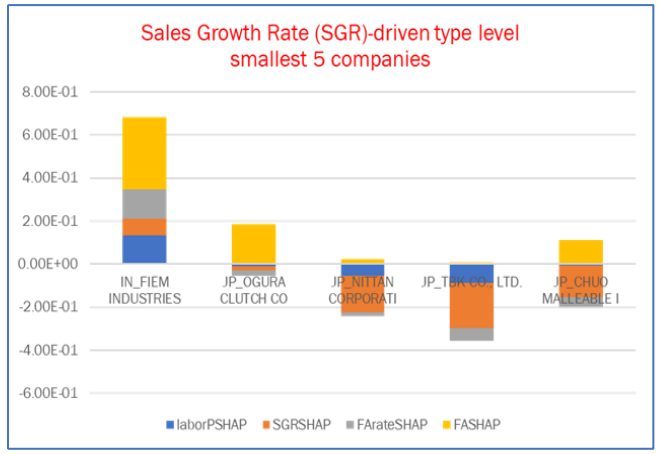

Figure 8. SHAP values of the SDG-driven type level smallest five companies.

Next, we will evaluate SHAP values of the SDG-driven type level smallest five companies (see Figure 8). In the three companies with positive SHAP values, FA_SHAP (Fixed Assets SHAP) is large and SGR_SHAP (Sales Growth Rate SHAP) is small, indicating that the companies are driven more by the size of their tangible assets rather than sales growth. For the two companies with negative SHAP 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 in these cases.

VI. EXISTING WORK

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

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 [18]-[23]. 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 true market information, while smaller ones reflect noise.
3. Singular Value Decomposition (SVD): SVD is applied to further clean the correlation matrix, focusing on the

Copyright (c) IARIA, 2025. ISBN: 978-1-68558-244-9

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.

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 [19]-[25].

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 [11, 13]. The two kinds of EigenVectors and the relationship among the three decomposed matrixes are clearly explained using visualization in [19, 20].

VII. Discussion

In this section, we will discuss the result. The objective of the analysis is a grouping of companies. The proposed SHAP_SVD method can extract the essence of the given SHAP value matrix. In the paper, the two extracted concepts were (1) Balanced (Well-balanced) type, and (2) Sales Growth Rate (SGR)-driven type. 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 6, the scattering plot of the companies by CompanyEigenVectors can uncover the individual companies' characteristics. The horizontal axis represents the Balanced (Well-balanced) type level. These higher-level companies can be divided by the vertical axis into two groups, which are 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 reflect each company's characteristics more accurately 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 Key Performance Indexes (KPIs) 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 important 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.

VIII. Conclusions

In this paper, we proposed the SHAP_SVD method, which combines SHAP values with SVD for regression analysis. The SHAP_SVD method enables the extraction of core concepts from the SHAP value matrix, providing a more accurate representation of each company's characteristics compared to using raw input data. In our case study, we identified two main concepts: (1) Balanced (Well-balanced) type and (2) SGR-driven type. By analyzing the SHAP values through SVD, we were able to group companies based on their unique feature contributions, revealing distinct business drivers.

This method offers deeper insights into corporate data by aligning the dimensionality reduction process with the specific characteristics of each company. Furthermore, the eigenvectors derived from SHAP_SVD can serve as a foundation for developing effective KPIs, helping businesses to identify and focus on the factors that most significantly influence their performance.

Acknowledgment

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

References

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

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

[3] 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.

[4] S. M. Lundberg, G. G. Erion, and S.-I. Lee, "Consistent individualized feature attribution for tree ensembles," *arXiv preprint arXiv:1802.03888*, 2018.

[5] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems*, vol. 30, 2017.

[6] S. M. Lundberg and S.-I. Lee, "Consistent feature attribution for tree ensembles," *arXiv preprint arXiv:1706.06060*, 2017.

[7] 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.

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

[9] A. Chaddad, J. Peng, J. Xu, and A. Bouridane, "Survey of explainable AI techniques in healthcare," *Sensors,* vol. 23, no. 2, p. 634, 2023.

[10] 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.

[11] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning* (no. 4). Springer, 2006.

[12] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*. Springer, 2009.

[13] S. Theodoridis and K. Koutroumbas, *Pattern recognition*. Elsevier, 2006.

[14] 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.

[15] 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), p. Decision Science Institute Best Paper Award.

[16] 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.

[17] XGBoostDevelopers. "XGBoost Decumentation (Revision 534c940a.)." https://xgboost.readthedocs.io/en/stable/ (accessed 2022/11/13.

[18] 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.

[19] 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: http://www.iaiai.org/journals/index.php/IEE/article/view/70.

[20] 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.

 ISBN: 978-1-68558-244-9