# Feature Importance Analysis in Global Manufacturing Industry

Kenji Yamaguchi

***Abstract*—SHAP is a measurement based on Shapley values and has been used widely in machine-learning regressions to interpret the feature importance. I conducted the feature importance analysis by the SHAP values in the global manufacturing industry. The target fields are automakers and electronic companies. I found the interesting attribute of Shapley values through the regression analysis. In general, the predictor variable values of companies forge no linear relationship to the target values such as a profit ratio. However, after making the SHAP values for each predictor, the scattering plot between the SHAP values and the target values clarifies the linear relationship between them. I verified the linear relationship on both automakers and electronic companies. The insight of the linearity is presented in this paper. Each company has a different behavioral structure specific to the company. The SHAP value extracts the company's behavioral structure through the characteristic function.**

**In addition, to make the regression results more precise and avoid effects by the multi-collinearity, I conducted a PCA (Principal Component Analysis). From the 3D scatter plot of the PCA of SHAP values, I verified the linear relationship as I expected and could identify the latent semantics of the PC1 and PC2 as a profitability related factor and an operation management relation factor.**

***Index Terms*—Shapley values, characteristic function, company performance measurement, machine learning, regression, global manufacturing.**

## I. Introduction

Machine learning based regressions improved the accuracy by using complex regression models, compared to the traditional multiple linear regression. However, it yields the problem which is lacks the interpretability. Despite widespread adoption, machine learning models remain mostly black boxes. Understanding the reasons behind predictions is, however, quite important if one plans to take action based on a prediction, or when choosing whether to deploy a new model [1].

For the interpretability, machine learning models can be interpreted through a number of new methods, such as MDI [2], MDA, LIME [3] and Shapley values. MDI (Mean-Decrease Impurity) evaluates the impurity as the entropy of the distribution of labels as the Gini index in a tree-based classification or regression [4]. A disadvantage of MDI is that a variable that appears to be significant for explanatory purposes (in-sample) may be irrelevant for forecasting purposes (out-of-sample) [1]. To solve the problem, Breiman has developed the mean-decreases accuracy (MDA) method [4].

Other approaches to estimate the importance of input predictors, are LIME, DeepLIFT [5], [6], and SHAP. Among them, in the machine learning regressions, currently SHAP by Lundberg is the most widely used as a unified framework for interpreting predictions [7]. The SHAP (SHapley Additive exPlanations) advantages are (1) SHAP is based the Shapley value which is theoretical results, showing that the Shapley value is the unique distribution solution with a set of desirable axioms such symmetry and linearity [8], [9], and (2) SHAP unifies a diverse set of the existing model explanation methods such as LIME an DeepLIFT [10].

The Shapley value is originally a solution concept in cooperative game theory [8], [9]. In a cooperative game, the Shapley value offers the unique distribution (among the players) of a total surplus generated by the coalition of all players. The SHAP values leverage these Shapley value properties. Shapley or SHAP values are widely used by machine learning models in various economic fields [11]-[13]. However, problems with Shapley-value-based explanations as feature importance are still left [14]. Ghorbani said, concerning Shapley values, as follows: *Drawing on the connections from economics, we believe the three properties we listed is a reasonable starting point. While our experiments demonstrate several desirable features of DATA SHAPLEY, we should interpret it with care [15]*. To incorporate the Shapley values with machine learning approaches, many researches have been still conducting evaluations [16]-[18].

This paper will illustrate the Shapley values' intrinsic meaning. The handling field is the global manufacturing industry. We have conducted many regressions to measure the performance levels of each company. The predictors of the regression are, for example, ROE (Return on Equity) and net sales, and we have evaluated the predictor importance, using SHAP values. Through the analyses, I found an interesting result that there was a linear relationship between the target values and the SHAP values. In most of all cases, there was initially no linear relationship between the target values and the raw predictor values. However, a linear relationship comes to appear between the target values and the SHAP values of predictors in both automakers' data and electronic industries' data. I had been thinking about the reason of the linearity appearance. Here in the paper, I shall illustrate the resultant linear relationships via concrete practical examples. I think that the linearity indicates the intrinsic potential of Shapley values.

So far as I know, there is no existing papers which explicitly described the linearity of SHAP values from the viewpoint of business and management. One paper had an

Manuscript received January 28, 2021; revised August 11, 2021.

K. Yamaguchi is with Science & Education Center (SEC), Ochanomizu University, Japan (e-mail: yamaguchi.kenji@ocha.ac.jp).

doi: 10.18178/ijtef.2022.13.2.719

explanation of the linear relationship as the principle behind Shapley regression [16] as follows: *Shapley values project unknown but learned functional forms (original input space) into a linear space (Shapley value space).* However, to the best of my knowledge, there is no other paper which described the linearity. Even in Lundberg's paper in 2018 [19], I cannot find descriptions of this linearity. As shown in [16], mathematically the linearity of Shapley values must have been proved already. However, I think that the details of the intrinsic meaning behind Shapley regression should be investigated in each application field.

The number of researches in the operations management field using Shapley-value-based evaluation is very limited. Although there are some papers using Shapley values in operations management fields, these papers had no description about the linearity I found [20], [21]. I would like to investigate the intrinsic meaning of the Shapley values in a global company performance analysis.

In the next section, I explain the Shapley values and SHAP by Lundberg. In Section III, I present my case study results. In Section IV, I evaluate the intrinsic meaning of the linearity extracted between the target values and SHAP values. Then I conclude the paper.

## II. Expectation Shapley Value

In this section, I shall show the formula of the original Shapley values and the SHAP approach which is an expanded version for the machine learning approaches. The explanations here are cited from [22] and Roth's explanations of the Shapley value [23].

Shapley values are a solution of how they should distribute the total profit in a cooperative game. There are $N$ players and do the cooperation together to obtain more profit than one of work separately. The given data is a characteristic function $v$ of a gain profit by on all subset $S$ of $N$ players. The interpretation of $v$ is that for any subset $S$ of $N$, $v(S)$ is the worth of the coalition, in terms of how much "utility" the members of $S$ should divide among themselves.

The function $v(X)$ is a characteristic function $v: 2^N \to R$ where $N$ is the number of players (and in our case study, $N$ is the number of predictor variables). The only restriction on $v$ was $v(S \cup T) \geq v(S) + v(T)$ which means that the worth of the coalition is equal to at least the worth of its parts acting separately.

The problem is given the following three assumptions [10]:

(1) Efficiency axiom: The sum of the Shapley values of all players equals the value of the total coalition, so that all the gain is distributed among the players. In my case study, the sum of the Shapley value of all predictors equals to the value of the all predictors' coalition in each company.

(2) Symmetry axiom: If players $i$ and $j$ are equivalent in the sense that $v(S \cup \{i\}) = v(S \cup \{j\})$ for every subset $S$ of $N$ which contains neither $i$ nor $j$, then the Shapley values for $i$ and $j$ are equivalent $\phi_i(v) = \phi_j(v)$. In my case study, it means that if the effects to the performance by the two predictors are the same, then the two predictors' Shapley values are the same.

(3) Additivity axiom: For two games $v$ and $w$, $\phi_i(v) + \phi_i(w) = \phi_i(v + w)$ for all $i$ in $N$, where the game $[v + w]$ is defined by $[v + w](S) = v(S) + w(S)$ for any coalition $S$. In my case study, $v$ and $w$ are corresponding to two company characteristics functions.

There the unique solution exists which meets the above three axioms. Shapley showed the formula of the Shapley value for player $i$ as follows [8]:

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!\,(|N| - |S| - 1)!}{|N|!} [\, v(S \cup \{i\}) - v(S)]$$

There, every subset $S$ of $N$ which does not contain $i$.

$$\sum_{S \subseteq F \setminus \{i\}} \frac{|S|!\,(|F| - |S| - 1)!}{|F|!} \left[ f_{S \cup \{i\}}\left(x_{S \cup \{i\}}\right) - f_S(x_S) \right]$$

where $F$ is a set of players, and $S$ is a subset of $F$ which does not include $i$-th player $S \subset F \setminus \{i\}$. $|F|!$ is the permutations of the number of $F$. The term $\left[ f_{S \cup \{i\}}\left(x_{S \cup \{i\}}\right) - f_S(x_S) \right]$ is player $i$'s marginal contribution to coalitions $S$; the function $f$ is calculated with the input of the player set $\left(x_{S \cup \{i\}}\right)$ and then the function $f$ is calculated with the input of the players set $(x_S)$. The difference is used as the key part of the formula.

The Shapley value can be interpreted as the expected marginal contribution of player $i$. Suppose that the players enter a room in some order and that all $|F|!$ orderings are equally likely [9].

The first $|S|$ players can precede $i$, and the number of permutations is $|S|!$. Then $(|F| - |S| - 1)!$ different orders in which the remaining $(|F| - |S| - 1)$ players can follow. Then $\frac{|S|!\ \ (|F|-|S|-1)!}{|F|!}$ expresses the probability of adding $i$-th player to the subset $S$.

An example is shown here: $F = \{X_1, X_2, X_3\}$ and $i = 1$ are given. In that case, $S$ are the following 4 subsets:

$$\left\{\{\ \ \}, \{X_2\}, \{X_3\}, \{X_2, X_3\}\right\}$$

The number of members are $|F| = 3$. Then the permutation is $|F|! = 6$.

- Case $S = \{\ \ \}$: Possible orders are (1) $X_1 \to X_2 \to X_3$ and (2) $X_1 \to X_3 \to X_2$. Then $\frac{|S|!(|F|-|S|-1)!}{|F|!} = \frac{2}{6}$.
- Case $S = \{X_2\}$: Possible orders are (1) $X_2 \to X_1 \to X_3$. Then $\frac{|S|!(|F|-|S|-1)!}{|F|!} = \frac{1}{6}$.
- Case $S = \{X_3\}$: Possible orders are (1) $X_3 \to X_1 \to X_2$. Then $\frac{|S|!(|F|-|S|-1)!}{|F|!} = \frac{1}{6}$.
- Case $S = \{X_2, X_3\}$: Possible orders are (1) $X_2 \to X_3 \to X_1$ and (2) $X_3 \to X_2 \to X_1$. Then $\frac{|S|!(|F|-|S|-1)!}{|F|!} = \frac{2}{6}$.

Then we consider the calculation method SHAP in a machine learning regression. We have to calculate each company characteristic function $v$ using the regression model $f(X)$.

When we try to find $v$, we have to fix every feature value to find the value of $v(x)$. So, we use the expected value (the average value) of the feature instead of the missing feature [7].

$$\phi_0 = E[f(X)]$$
$$\phi_1 = E[f(X)|X_1 = x_1] - \phi_0$$
$$\phi_2 = E[f(X)|X_1 = x_1, X_2 = x_2] - \phi_1$$
$$\phi_3 = E[f(X)|X_1 = x_1, X_2 = x_2, X_3 = x_3] - \phi_2$$

The above four expressions are cited from Fig. 1 in [7] explain how to get from the base value $E[f(X)]$ that would be predicted if we did not know any features to the current output $f(x)$. The SHAP values arise from averaging the $\phi_i$ values across all possible orderings.

The SHAP value $\phi_i$ may be a negative value. Therefore, to calculate the importance of a feature, the average of all the absolute values $\phi_i$ is used. The importance of a feature $i$ is calculated as $\frac{1}{N}\sum_{k=1}^{N}|\phi_i^k|$ where $N$ is the sample size.

The original Shapley value is the solution of how to divide the total profit to every player and the expected value of each player's marginal contribution to the total profit. The Shapley value is calculated under the condition that all possible orderings are equally likely. When we use the Shapley value in a regression model, a player is a predictor feature and the profit is the target value. The SHAP value can provide us with the marginal contribution of each predictor feature.

Then, let us consider the measurement of relative importance of a predictor in a regression analysis. In the tree based regression, for example, Random Forest, the relative importance of a predictor was first defined as the sum of squared improvements over all internal nodes for which the feature was chosen as the splitting feature [24]. Compared to this definition, the SHAP value is the bottom-up procedure from each company's characteristics. That is the advantage of the SHAP values. The previous method which counts levels of purification improvement by splitting the area, being ignored each company's characteristics; in other word, the previous method ignores which companies are split to which areas. What counts on there was just the improvement of the purification level of split areas. Another advantage of the SHAP is availability to any regression models such as XGBoost or Random Forest. SHAP is available, so far as we can have the regression model.

Let us think about this case study. The regression model $f$ is calculated from all company data. Using the model, we make a characteristic function $v$ for each company, $v: 2^n \rightarrow R$ where $n$ is the number of predictors. The resultant SHAP value reflects how a company behaves for the target value. The $n$ feature values are related to one another within the company. The SHAP value can then be interpreted as an expected utility of each predictor for the target. However, the function $v$ is made based on the regression model *f*. The calculation of the regression model *f* requires all *n* feature values; We cannot calculate characteristic values by *f*, like $v: 2^n \rightarrow R$. To solve the problem, we have to use the average value of all companies in *f*, instead of the missing feature.

Then it becomes important to select the domain of the companies. We must select a set of company data from the same industry field such as an automaker and a machinery. Generally speaking, when we evaluate the performance of a company, we will see whether the company's performance is greater than the average level of the industry field or under the level, compared to the same industry companies. Because we use the average level, we have to be careful when we select the set of companies. This is because we would like to investigate a company's behavior through the comparison to others in the same industry.

## III. Case Study

In the section, I shall present the linearity between target values and SHAP values, using a concrete practical example.

The case study that I use here is a regression analysis of the stock price data of leading global manufacturers. The industry fields are (1) automakers such as Toyota and (2) electronic companies such as Sony. The data we used is stock price data. Its period or so is shown in Table I.

TABLE I: Manufacturing Companies' Data

| Industry Field | Automakers | Electronics Companies |
|---|---|---|
| Stock Price Data Period | March 1, 2018 to April 20, 2018 | March 1, 2018 to April 20, 2018 |
| Regression Data Period | 2013 to 2017 | 2013 to 2017 |
| Number of Companies | (total) 54 | (total) 100 |
| Germany | 7 | 5 |
| Japan | 34 | 45 |
| US | 13 | 50 |

I conducted the regression on the data. The target variable of the regression is the ratio of the initial stock price and the final stock price. In the period, the stock prices were decreased severely owing to the President Trump's remarks concerning the US-China trade friction. I made the ratio as the target variable, so that I wanted to investigate the recovery power of companies. The stock price reflects the investors' speculation for a company and evaluation of a company. Generally speaking, a company which investors evaluate as a high-performance company has a high recovery power. The features/predictor variables are shown in Table II.

From a viewpoint of business and management, these six predictors are divided to (1) the profitability related ones [PBT, ROA and ROE] and (2) the operations management related performances [IVR and FAR), and (3) the sales growth rate [SGR]. High inventory ratio and high tangible fixed assets ratio mean "slow" and "heavy" supply chains. SGR is introduced to stand for "growth potentiality".

Details of the regression results of the automakers are described in [25]. In this paper, I shall focus on the SHAP interpretation.

TABLE II: Six Predictor Variables of the Regression

| Predictor | Explanation |
|---|---|
| 1) SGR | Sales growth ratio (the geometric mean of the five years) |
| 2) PBT | Profit before tax ratio (over sales) |
| 3) ROA | Returns on assets |
| 4) ROE | Returns on equity |
| 5) IVR | Inventory ratio (over sales) |
| 6) FAR | Tangible fixed assets ratio (over sales) |

The database of the six predictors was ORBIS by Bureau Van Dijk. I would like to investigate the long term effects. Therefore, I set the predictor variables of the regression to be averages during the past five years (2013- 2017).

The stock price data were also taken from the ORBIS database. I conducted the regression analysis by XGBoost method. The used environment is the python environment with the scikit-learn libraries (https://scikit-learn.org/stable/index.html) [26], [27]. The SHAP library I used is presented by Lundberg as GitHub (https://github.com/slundberg/shap) [7].

The resultant $R^2$ of XGBoost regression was 1.0 in both the automakers and the electronics companies. This is so called "over-fitting" which should be avoided to make prediction. In this analysis, however, my objective is to illuminate latent relationship structures between the stock price recovery (target) and the six performance measures (predictors). For the purpose, I used all data as the training data without test data, so that I can see the whole companies' behaviors.

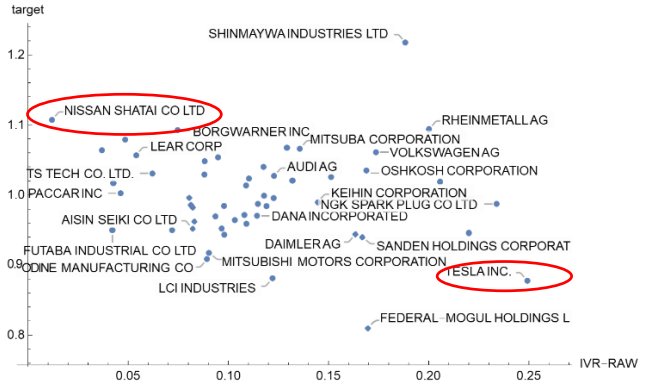


Fig. 1(a). Plot between target values and raw predictor IVR in automakers (The correlation coefficient is -0.108318).

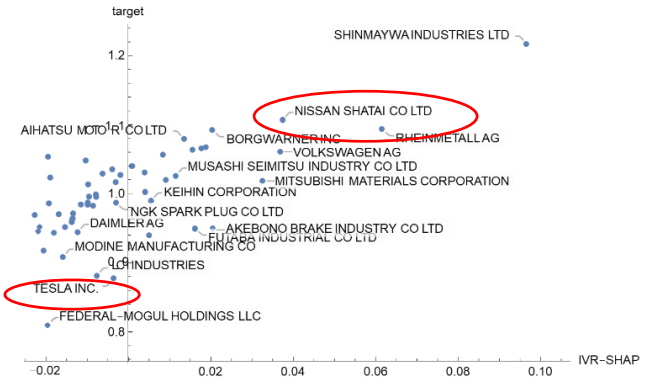


Fig. 1(b). Plot between target values and SHAP IVR in automakers (The correlation coefficient is 0.586592).

In Fig. 1, I show six comparisons between (a) the plot between raw predictor IVR values and the target values, and (b) the plot between the IVR's SHAP values and the target values. As shown in Fig. 1(a), there is no simple relationship between the IVR and the target values. The correlation coefficient is also -0.1. On the other hand, using the IVR SHAP values, the linearity appeared (See Fig. 1(b)). The correlation coefficient becomes 0.59. I found that a small IVR value contributed more to the target value (See SHAP values in Fig. 1(b)). From the theory, the relationship between the Shapley values and the target values is always positive, because the original Shapley value definition is a marginal effect of a predictor (player) to the target value (profitability); The original assumption is that the cooperation (coalition) makes the worth better than the worth of its parts acting separately.

This IVR result supports that a company with a low inventory ratio in general can recover its stock prices quickly. In Fig. 1(a), we can see TESLA and NISSAN SHATAI. The inventory ratio of TESLA is larger than that of NISSAN SHATAI; A heavy inventory is not good from the viewpoint of operations management. After the SHAP calculation, in Fig. 1(b), I can see that the small inventory ratio by NISSAN SHATAI still more positively contributes to the increase of the target value, compared to the high inventory ratio by TESLA.

Then, from a comparison between Fig. 1(c) and Fig. 1(d), I found that a company with a low sales growth ratio recovered its stock prices quickly as well as the IVR comparison. Contrary to my expectations, concerning SGR (sales growth rate), a recovery of a company with a rapid growth rate was slow, although the relationship is not so strong.

In Table III, the total comparison concerning the automakers' results are shown. The left row shows the predictors' correlation coefficients. In the right row, the six correlation coefficients of SHAP values which are all positive.

Then let me show the result of the electronic companies. First, let me show the improvement concerning IVR (See Fig. 2). Although there is no relationship between the predictor IVR and the target, the linear relationship with the correlation coefficient 0.81 clearly appeared by the SHAP values. In other predictor comparisons, we can see the same kind of improvement of correlation changes toward the linearity, as shown in Table IV. Especially the correlation of IVR SHAP values is very high 0.81.

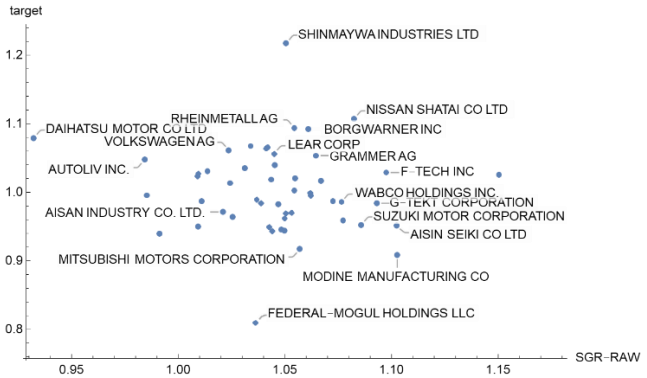


Fig. 1(c). Plot between target values and raw predictor SGR in automakers (correlation coefficient is -0.271465).

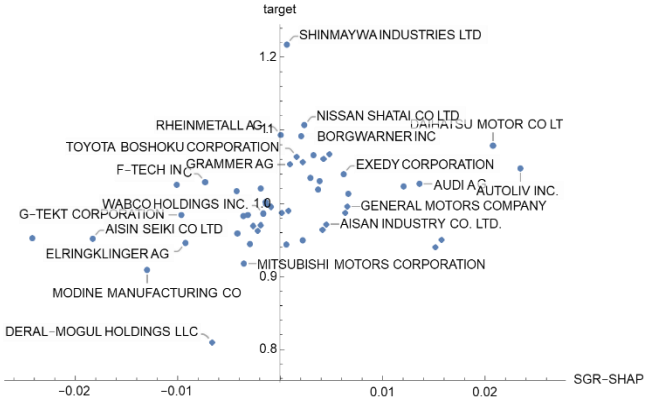


Fig. 1(d). Plot between target values and SHAP SGR in automakers (correlation coefficient is 0.362226).

TABLE III: AUTOMAKERS' IMPROVEMENT OF CORRELATION COEFFICIENTS

| Predictor | Raw | SHAP |
|---|---|---|
| 1) SGR | -0.271465 | 0.362226 |
| 2) PBT | 0.364203 | 0.737775 |
| 3) ROA | 0.340880 | 0.317088 |
| 4) ROE | 0.509682 | 0.769282 |
| 5) IVR | -0.108318 | 0.586592 |
| 6) FAR | -0.207072 | 0.660774 |

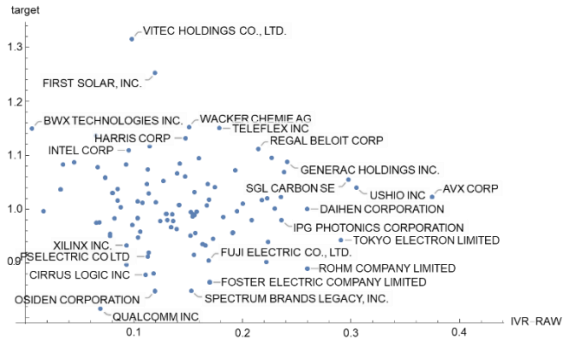


Fig. 2(a). Plot between target values and raw predictor IVR in electronic companies (The correlation coefficient is -0.105506).

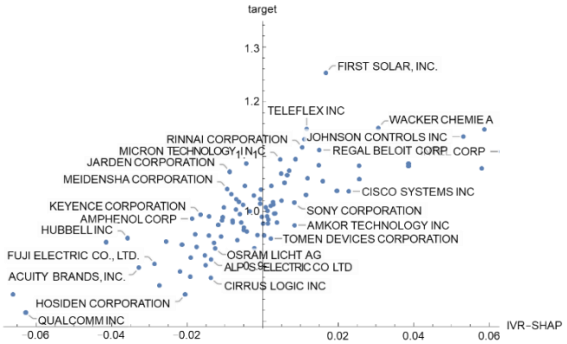


Fig. 2(b). Plot between target values and SHAP IVR in electronic companies (The correlation coefficient is 0.806714).

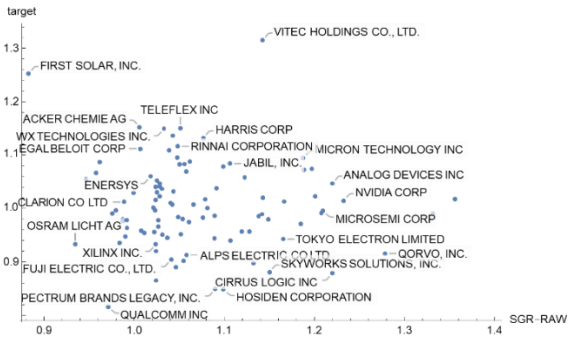


Fig. 2(c). Plot between target values and raw predictor SGR in electronic companies (The correlation coefficient is -0.069889).

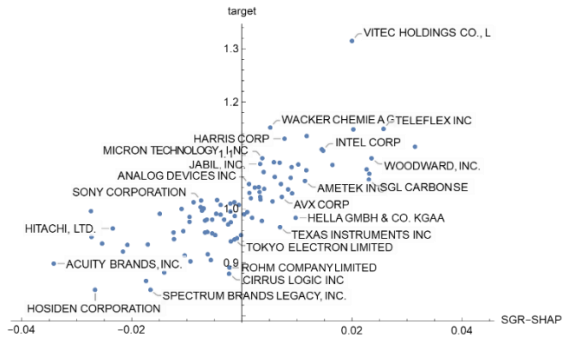


Fig. 2(d). Plot between target values and SHAP SGR in electronic companies (The correlation coefficient is 0.694865).

TABLE IV: ELECTRONICS COMPANIES' IMPROVEMENT OF CORRELATION COEFFICIENTS

| Predictor | Raw | SHAP |
|---|---|---|
| 1) SGR | -0.069889 | 0.694865 |
| 2) PBT | -0.057337 | 0.712772 |
| 3) ROA | -0.077992 | 0.620517 |
| 4) ROE | 0.072975 | 0.699969 |
| 5) IVR | -0.105506 | 0.806714 |
| 6) FAR | 0.109430 | 0.680907 |

## IV. SHAP VALUE AS PERFORMANCE MEASUREMENT

In general, when we compare rival companies in the same industry field, we use the predictor raw value. For example, a company with ROA (Return On Assets) over 5% is considered to be a high performance company. However, I think that the ranking of this raw predictor data hardly shows the order of the high performance of companies. The raw predictor values lack the viewpoint of companies' characteristics. If we know the company behavior, we should utilize the company's characteristic model information. I noticed this important thing, from the comparisons of correlation coefficients shown in Fig. 1 and Fig. 2.

Let me use an analogy for the explanation. In a human body, a systolic blood pressure 135 mmHg is considered to be unhealthy. However, even with the value 135 mmHg, some persons are healthy. Whether healthy or not depends on the person's physical characteristics. Let me back to this case study. For example, suppose that there are two companies X and Y; They are company X with ROA 4% and company Y with ROA 12%. The ROA values do not indicate contribution to the target values if in company X, the contribution by ROA (4%) might be much higher than ones of other five predictors and if in company Y, the contribution by ROA (12%) might be much smaller than ones of other five predictors. The absolute ROA values cannot reflect the company's internal behavior or characteristics.

On the other hand, in the evaluation by SHAP in the regression, the SHAP can reflect the characteristics of the company. In the process, I first conduct a regression to obtain the regression model. Next, using the regression model, I calculate the company X's characteristic function *v*. Finally, using the characteristic function, I shall calculate the SHAP value of each predictor concerning company X. The SHAP of ROA indicates how much effective the ROA value is in order to increase the target value within the company X. On the effectiveness evaluation, the company X's behavior or structural characteristics is reflected.

For example, suppose that company Y's IVR is much higher than the average. The higher IVR in general leads to a delay of the recovery after the large decline of stock prices. However, within company Y, higher IVR might play a good role for a quick recovery. If so, the SHAP of IVR within company Y would become positively large.

## V. EVALUATION BY PCA

In the section, I conduct a PCA (Principal Component Analysis) to remove the multi-collinearity among the predictors.

The problem by multi-collinearity is called substitution

effects in machine learning methods. Substitution effects can bias the results from feature importance methods. In the case of MDI (Mean decrease impurity), the importance of two identical features will be halved, if they randomly chosen with equal probability [1]. In general, to solve the multi-collinearity problem, we apply PCA. The orthogonalization of features by PCA can remove the multi-collinearity [1].

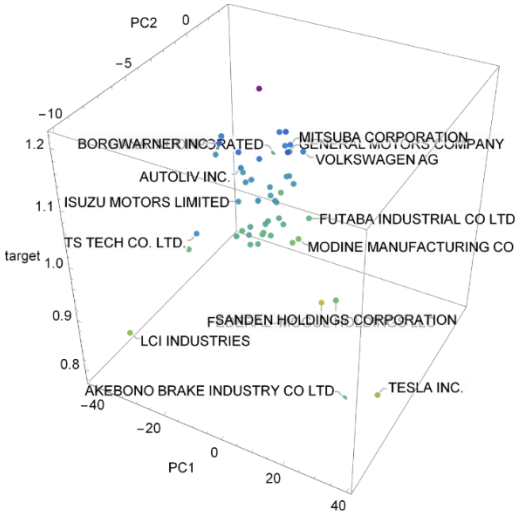


Fig. 3. Raw Predictor: 3D scatter plot between PC1, PC2 and target values in automakers.

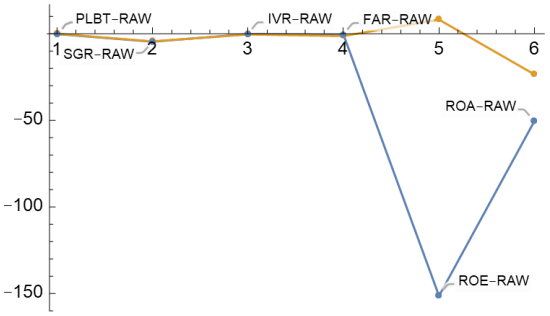


Fig. 4. Raw Predictor: Eigenvectors of PC1 and PC2 in automakers. The first eigenvector is in blue.

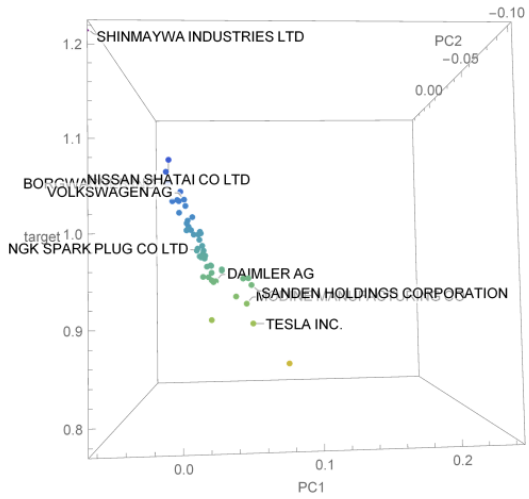


Fig. 5. SHAP: 3D scatter plot among PC1, PC2, and target values in automakers.

First, I conducted a PCA using the six raw predictor values of the automakers. Then I conducted the dimensional reduction from 6 dimensions to 2 dimensions using the first and second Principal Components. In Fig. 3, I showed the result of the 3-dimensional plot with the axes of PC1 (Principal Component first) and PC2, and the target values. As shown there, I found no linear plane among them. The corresponding first and second eigenvectors are shown in Fig. 4. The element values are as follows: PC1 {-0.485585, -4.87069, -0.472493, -0.936612, -151.265, -50.5595}, and PC2 {-0.161258, -4.51894, -0.503542, -1.44312, 8.06248, -23.6532}. The element values are plotted in Fig. 4. The eigenvectors seem to be meaningless.

Then I conducted a PCA of the six SHAP values. From the scatter plot by the SHAP values' PC1, PC2, and the target values, a clear linear relationship appeared (See Fig. 5). This is a very reasonable consequence, because each predictor SHAP values have a linear relationship to the target values which was shown in the previous section.

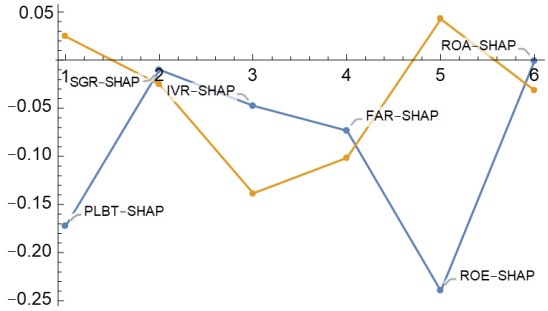


Fig. 6. SHAP: Eigenvectors of PC1 and PC2 in automakers.

Fig. 6 shows the first and second eigenvectors. In the first eigenvector which is in blue, ROE(5) and PLBT(1) are the main factors, because their absolute values are large. In the second eigenvector, IVR(3) and FAR(4) are the main factors. Then, the first PC can be interpreted as a profitability relted factor. The second PC can be interpreted as an operation management related factor. I can obtain clearly the latent meanings of the PCs.

Then, let's see the electronics companies' results. Fig.7 shows the predictors' PCA results. Like the automakers' one, there is no linear relationship there. The eigenvectors also have no meanings as shown in Fig. 8. On the other hand, using SHAP values, the linear relationship clearly appeared (See Fig. 9). Concerning the eigenvectors of the first eigenvector, in the negative side there are mainly IVR, FAR, and ROE. In the second eigenvector, in the positive side there are mainly IVR, FAR, and ROE (See Fig. 10).

In general, investors are firstly interested in the profitability features such as PLBT, ROA, and ROE. The IVR and FAR are likely to be forgotten. However, I think that the contribution by IVR and FAR are great and should be analyzed precisely. The SHAP PCA results also showed that IVR and FAR contributed a lot as the main factors. To extract the effects by the operation management contribution, a PCA by raw predictor values are not useful and the PCA by SHAP values are effective.

SHAP values are measurement of predictors, taking into account of company characteristics. I would like to propose that the SHAP values should be used as a company performance evaluation as follows: First, in the same industry group, we conduct the regression and evaluate the SHAP values for each company. In each predictor, the average and standard deviation of SHAP values are calculated. Using the

values, the deviation score of each company can be found. In Fig. 11, sample radar charts of two companies are illustrated using the deviation scores. The blue company's scores are over 50 in many predictors, which means the company is the higher performance company than the average level. Compared with the orange company, in every aspect except ROE, the orange company scores are higher than those of the blue company. Although the blue company is well-balanced, the orange company is a higher performance company. In addition, I think that the orange company can increase the ROE values still more that is the stretch point of the company.

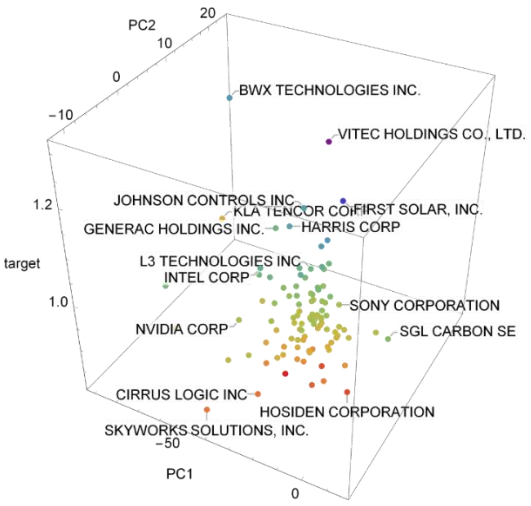


Fig. 7. Raw Predictor: 3D scatter plot between PC1, PC2 and target values in electronic companies.

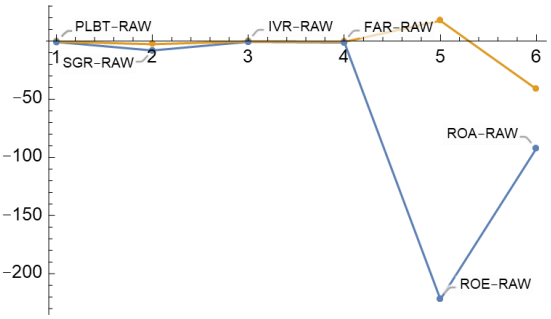


Fig. 8. Raw Predictor: eigenvectors of PC1 and PC2 in electronic companies.

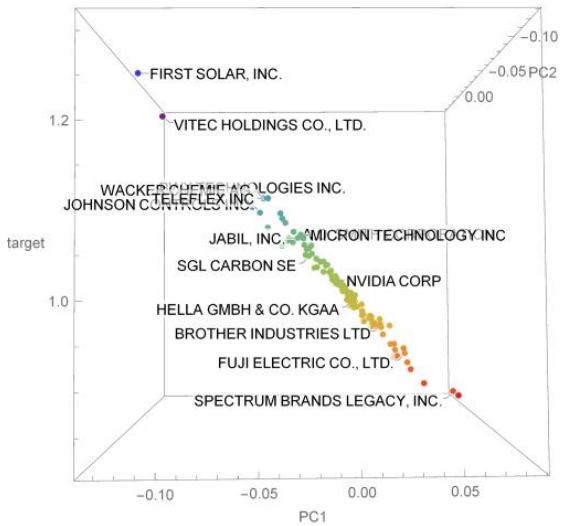


Fig. 9. SHAP: 3D scatter plot between PC1, PC2 and target values in electronic companies.

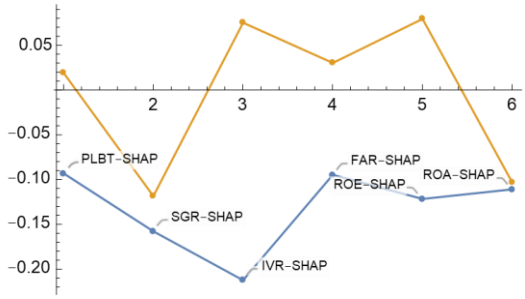


Fig. 10. SHAP: eigenvectors of PC1 and PC2 in electronic companies.

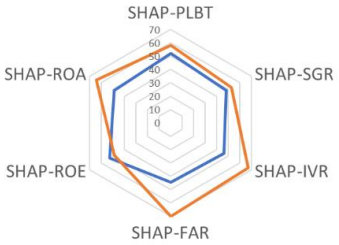


Fig. 11. Sample of a SHAP based radar chart.

## VI. Conclusion

In the paper, I described the feature importance analysis of global manufacturing companies. The industry fields are (1) automakers and (2) electronics companies. In general, we had not found the linear relationship between the raw predictor values and the target values in the regression. On the other hand, using the SHAP values, I found the linear relationship between the SHAP values and the target values in each predictor's regression.

The original point of the paper is that I showed the linear relationship between the target and the SHAP values in both industry fields. Secondly, I offered a new interpretation of the SHAP's intrinsic meaning. A SHAP value is calculated based on each company's characteristics, and the SHAP value reflects the company's behavioral structure, which is the intrinsic meaning of SHAP values in a company performance evaluation. Therefore, when we evaluate the company performance, SHAP is better that the raw predictor values.

A SHAP value reflects how a company behaves for the target value and the N predictor values are related to one another within the company. Therefore, a linear relationship appears between the target values and the predictor SHAP values, even when there was no linear relationship between the target values and the raw predictor values.

The reason why we cannot find the linearity in the regression by the raw predictor values has no relationship to the multi-collinearity problem. To show that, I conducted a PCA by predictor values and using the PC1, PC2, and the target values, I illustrated the 3D scatter plot. There is no linearity. On the other hand, from the PCA of SHAP values, I could extract latent semantics of the PC1 and PC2 such as a profitability related factor and an operation management relation factor.

It is interesting that after reflecting the company's characteristics by the SHAP approach, the regression result showed the linear relationship in all predictors and the target values in this case study. This visualization made us understand the intrinsic meaning and potential of SHAP values.

## Conflict of Interest

The research was partly supported by the fund of Gakushuin University Computer Centre project 2020 and by JSPS KAKENHI Grant Number 20H01537.

## Acknowledgment

I am deeply grateful to Professor Yukari Shirota for thoughtful guidance throughout in this analysis.

## References

[1] M. L. de Prado, *Machine Learning for Asset Managers*, Cambridge University Press, 2020.
[2] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Regression Trees*, CRC press, 1984.
[3] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you?" Explaining the predictions of any classifier," in *Proc. the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 1135-1144.
[4] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5-32, 2001.
[5] A. Shrikumar, P. Greenside, A. Shcherbina, and A. Kundaje, "Not just a black box: Learning important features through propagating activation differences," *arXiv preprint arXiv:1605.01713*, 2016.
[6] A. Shrikumar, P. Greenside, and A. Kundaje, "Learning important features through propagating activation differences," *arXiv preprint arXiv:1704.02685*, 2017.
[7] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in Neural Information Processing Systems*, pp. 4765-4774, 2017.
[8] L. S. Shapley, *Notes on the n-Person Game—II: The Value of an n-Person Game*, 1951.
[9] A. E. Roth, "Introduction to the Shapley value," *The Shapley Value*, pp. 1-27, 1988.
[10] S. Lundberg and S.-I. Lee, "An unexpected unity among methods for interpreting model predictions," *arXiv preprint arXiv:1611.07478*, 2016.
[11] H. Shalit, "The Shapley value of regression portfolios," *Journal of Asset Management*, pp. 1-7, 2020.
[12] M. V. Garcı́a and J. L. Aznarte, "Shapley additive explanations for NO2 forecasting," *Ecological Informatics*, vol. 56, p. 101039, 2020.
[13] F. Dong, B. Yu, Y. Pan, and Y. Hua, "What contributes to the regional inequality of haze pollution in C hina? Evidence from quantile regression and Shapley value decomposition," *Environmental Science and Pollution Research*, pp. 1-16, 2020.
[14] I. E. Kumar, S. Venkatasubramanian, C. Scheidegger, and S. Friedler, "Problems with Shapley-value-based explanations as feature importance measures," *arXiv preprint arXiv:2002.11097*, 2020.
[15] A. Ghorbani and J. Zou, "Data shapley: Equitable valuation of data for machine learning," *arXiv preprint arXiv:1904.02868*, 2019.
[16] A. Joseph, "Shapley regressions: A framework for statistical inference on machine learning models," *arXiv preprint arXiv:1903.04209*, 2019.
[17] K. Aas, M. Jullum, and A. Løland, "Explaining individual predictions when features are dependent: More accurate approximations to Shapley values," *arXiv preprint arXiv:1903.10464*, 2019.
[18] J. Chen, L. Song, M. J. Wainwright, and M. I. Jordan, "L-shapley and c-shapley: Efficient model interpretation for structured data," *arXiv preprint arXiv:1808.02610*, 2018.
[19] S. M. Lundberg, G. G. Erion, and S.-I. Lee, "Consistent individualized feature attribution for tree ensembles," *arXiv preprint arXiv:1802.03888*, 2018.
[20] E. A. Antipov and E. B. Pokryshevskaya, "Interpretable machine learning for demand modeling with high-dimensional data using Gradient Boosting Machines and Shapley values," *Journal of Revenue and Pricing Management*, pp. 1-10, 2020.
[21] M. Kraus, S. Feuerriegel, and A. Oztekin, "Deep learning in business analytics and operations research: Models, applications and managerial implications," *European Journal of Operational Research*, vol. 281, no. 3, pp. 628-641, 2020.
[22] S. M. Lundberg, G. G. Erion, and S.-I. Lee, "Consistent individualized feature attribution for tree ensembles," *arXiv:1802.03888*, 2018.
[23] A. E. Roth, *The Shapley Value: Essays in Honor of Lloyd S. Shapley*, Cambridge University Press, 1988.
[24] L. Breiman and R. Ihaka, *Nonlinear Discriminant Analysis via Scaling and ACE*, Department of Statistics, University of California, 1984.
[25] K. Yamaguchi, Y. Shirota, and M. Morita, "Effects of political risks on stock prices under global operations: A case study of US-China trade friction," in *Proc. EurOMA Conference*, University of Warwick, United Kingdom, 2020, pp. 582-591.
[26] D. Nedal and P. Morgan, *Introduction to Machine Learning with Python: A Guide for Beginners in Data Science*, 2018.
[27] J. Anderson, *Hands On Machine Learning with Python*, CreateSpace Independent Publishing Platform, 2018.



**Kenji Yamaguchi** was specially appointed as a lecturer in Science & Education Center (SEC), Ochanomizu University. His research fields are in stock price data, data science, science education, educational technology, learning support system and programming education. He has worked in the following workplaces: Computer Centre, Gakushuin University from 2010 to 2014; IT Center, Ochanomizu University from 2014 to 2017; Ochanomizu University Senior High School from 2017. He is a member of European Operations Management Association (EurOMA), Japanese Operations Management and Strategy Association (JOMSA), Japan Society of Business Mathematics, and Information Processing Society of Japan.