A Study on the Use of Artificial Intelligence within Government Pension Investment Fund's Investment Management Practices (Summary Report)

Takahiro Sasaki
Hiroo Koizumi
Takao Tajiri
Hiroaki Kitano

Sony Computer Science Laboratories, Inc.

March 2018

This report should be cited as: Sasaki, T., Koizumi, H., Tajiri, T., Kitano, H. (2018). A Study on the Use of Artificial Intelligence within Government Pension Investment Fund's Investment Management Practices (Summary Report). Tokyo, Japan: Government Pension Investment Fund.

Abstract

The proper selection and monitoring of fund managers is one of the most important tasks for Government Pension Investment Fund (GPIF). Its current approach, which depends on the track records and qualitative explanation of candidates and commissioned fund managers, can be significantly improved with the use of Artificial Intelligence (AI). This would lead to better management of GPIF's assets, which amount to over 150 trillion yen. GPIF expressed their interest in using AI technologies to improve overall business practices, especially on their core competences. Thus, this commissioned study focused on the investigation of: (1) The possibility and implications of applying AI technologies to the long-term management of pension assets, and (2) The impact of AI technology on the business models of asset management companies, especially if GPIF establishes its AI capability to select and monitor fund managers.

One of GPIF's core activities is the development and maintenance of a "manager structure". "Manager structure" refers to the structure of managing organizations (managers), and the associated allocations and re-allocations of assets to be managed. Within the process of developing and maintaining the manager structure, it is necessary to make various decisions, such as defining the characteristics of each manager, defining their management behavior with respect to various aspects of the economic background, and determining whether their behavior is consistent with the policy declared to GPIF.

The fundamental questions for this study are, "Is there any chance AI could be used within manager structure development and maintenance processes?", and "In which part of the process can we use AI most effectively?" A joint team formed between GPIF and Sony CSL have gone through GPIF's manager structure development and maintenance processes in-depth, and agreed to develop a proof-of-concept prototype system to test the principle of using deep learning to detect the investment style of managers from trading behavior data (trading items, timing, volume, unrealized gain and loss, etc) collected daily by GPIF. The system is composed of a series of "detector arrays", each reacting to the specific investment style of each manager. A detector array is a set of neural networks that are trained beforehand with the data generated by virtual fund managers, that each faithfully execute one of the typical investment styles. The system provides an N-dimension vector representing a mixture of investment styles of the manager at given point in time. A blind test of style detection using actual trading behavior data for 16 fund managers demonstrated that the system can properly detect the styles and drifts of each fund manager. In addition, the system's visualization capabilities were proven to be effective in identifying the spontaneous convergence of trading behaviors where most funding managers happen to trade similar items.

Results from the proof-of-concept pilot system indicate that, with the introduction of such a system, GPIF should be able to conduct more prudent and data-driven selection and monitoring of fund managers. In addition, this may foster more constructive and in-depth dialog between GPIF and fund managers, which will improve the robustness and performance of investment practices at GPIF in the long run.

Chapter 1

Issues within GPIF's Manager Structure Development and Maintenance Practices

GPIF, reflecting its high outsourcing ratio compared to pension funds in other countries, has a strong awareness of issues regarding the outsourcing cost of active investment management, and the improvement of the manager selection process, including the strengthening of knowledge and skills of personnel involved. GPIF's annual report for the 2016 fiscal year[1] argues that the active return, an excess earning of active funds, has been disappointing in the last 10 years, with domestic bonds at -0.12%, domestic stocks at -0.29%, foreign bonds at +0.64% and foreign stocks at -0.70%. Except for foreign bonds, it has not been possible to attain alpha (active return). On the other hand, payments to asset management companies and asset managers amounted to 9.9 billion yen for domestic bonds, 13.7 billion yen for domestic stocks, 12.8 billion yen for foreign bonds, and 34.5 billion yen for foreign stocks (a cumulative total for three years). They are not negligible and make no sense. This resulted in the recent contractual policy of GPIF linking management fees with active return. However, it does not solve all the problems within the manager structure practices.

Given these circumstances, discussions concerning the selection of active managers took place in a series of meetings of Board of Governors within GPIF. Some of the concerns raised in these meetings, held in November and December of 2017[2, 3], are on the issues of, "There were comparisons of current funds and past contracted funds, but it was not possible to make comparisons with other, potentially better funds," "The approach of selecting a better fund from the point of view of the whole universe is desirable," and "There is a concern that the qualitative evaluation might be criticized as having a degree of arbitrariness and lack of objectivity." Furthermore, interviews with the person in charge of GPIF's management revealed that the evaluation and selection of active managers is implemented under very strong constraints by a small number of internal experts.

GPIF clearly recognized the need to fully modernize the process, possibly using data science and AI to rectify the problems within the current process. In addition, GPIF wishes to investigate the possible impact of new types of funds that utilize the latest AI technologies, such as Two Sigma[4, 5], Renaissance Technologies[6, 7], Investifai[8, 9, 10], and so on. The question here is whether those funds can be evaluated by applying existing methods, or whether new methodologies are required.

The objective of this research is clearly defined; it agrees to provide a proof-of-concept AI system that, with full deployment, enables GPIF to select and monitor fund managers based on data and stringent analysis of trading behavior data, and to foster productive and insightful dialog with fund managers. The introduction of such a system will contribute significantly to the improvement of its investment practice, while managing risk properly.

Chapter 2

The Possible use of AI Technologies within Manager Structure Development and Maintenance

Characterizing the investment behavior of fund managers using trading behavior data is one of the most critical technical challenges involved in the improvement of manager structure development and maintenance practices. Data-driven characterization of fund managers enables GPIF to make well-informed decisions regarding fund manager selection, and improve the quality of dialog with fund managers under an active contract. The proof-of-concept system is based on a deep learning neural network designed to detect investment styles from trading behavior data.

2.1 Visualizing the Diversity of Funds

Each fund has its own investment style that is founded on basic investment philosophy, strategy, and methods. The differences in investment style, by which each fund is characterized, ultimately appear in the daily trading behaviors. Visualization of trading behavior data covers some of the characteristics of investment style.

The trading behavior visualization for a number of funds that invest in domestic equities over the period from April 1, 2014 to March 30, 2018 is presented in Figure 2.1. In each chart shown in the figure, the individual stocks are arranged on the horizontal axis from left to right in descending order of market capitalization at the end of March 2018. The time in days is represented on the vertical axis. The blue dots indicate a "buy" trade for the corresponding stock on the horizontal axis and day on the vertical axis. The red dots indicate a "sell" trade and the uncolored symbol indicates that no trade was made.

For example, the right side of the chart for *fund "B"* is empty, because the fund deals mainly in large-capital stocks, whereas the chart for *fund "P"* is conversely empty on the left side because the fund deals mainly in small capitalization stocks. The chart for *fund "L"* exhibits the characteristics of a quantitative style, in which investment targets are bought and sold mechanically within a wide range based on mathematically determined decisions. An opposite approach can be seen in the chart for *fund "O"*, which adopts a discretionary method in which stocks are selected manually on the basis of careful investigation. From the chart, we can see that trades are performed in a narrow range that includes only a small number of stocks.

As we see, the charts that visualize trading behaviors reflect each distinctive investment style. In some charts, we can spot specific periods of relatively intensive trading activities, which are shown as dense patterns of dots. It turned out that the fund had been conducting a huge amount of trading operations, due to additional capital for investment or a large-scale re-allocation of assets.

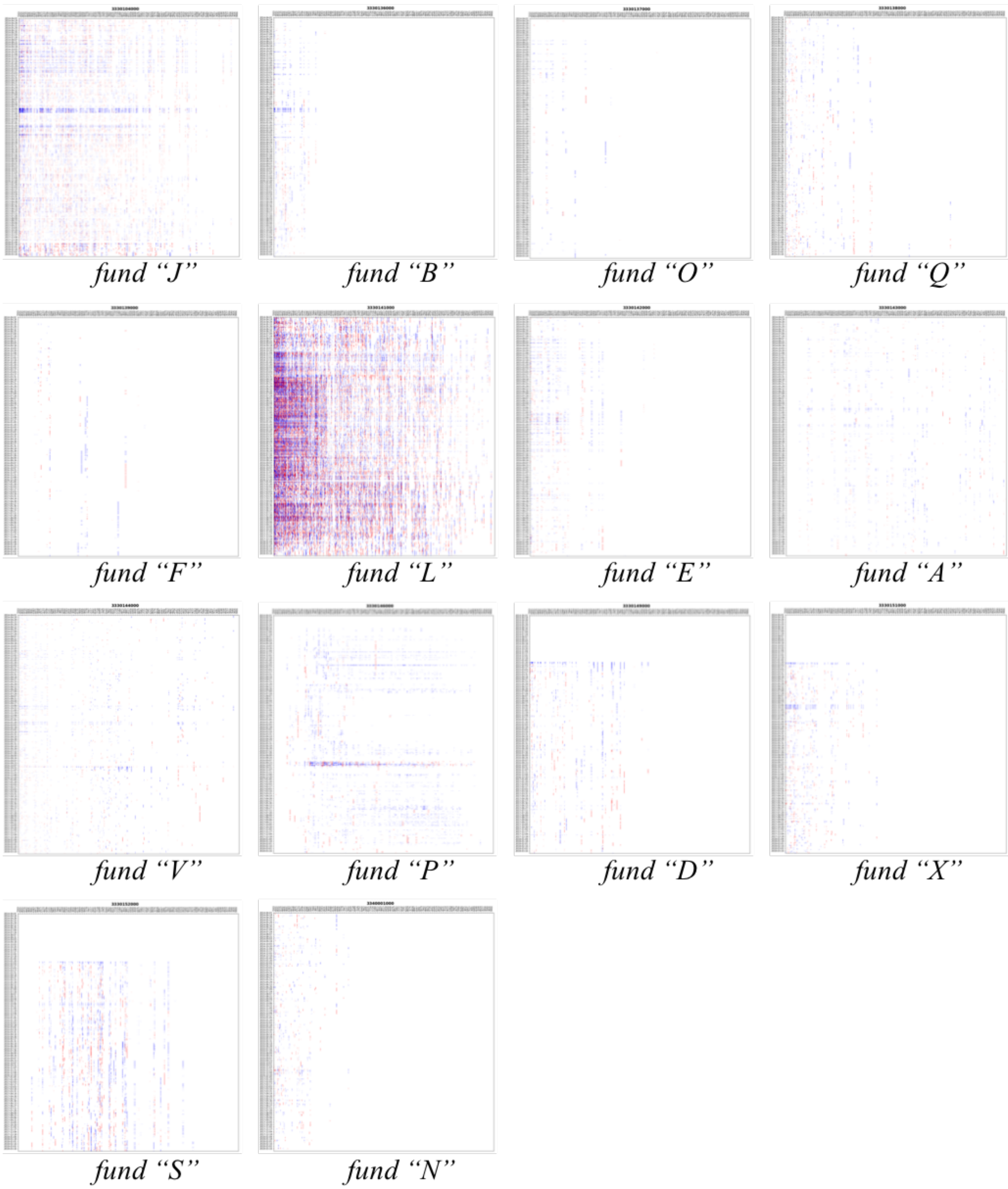


Figure 2.1: A visualization of fund manager trading behavior

2.2 Detecting "Style" using Deep Learning

To evaluate investment styles in more quantitative detail, we applied deep learning technology to devise an investment style analyzer, which we named a "Style Detector Array" system, and used it to analyze a number of domestic equity funds. Because our objective was to first test the operating principle, we restricted the universe to 100 stocks that include mainly large market capital issues. The Python programming language was used for implementation with the core libraries of deep learning such as tensorflow[11], Keras[12], and complementary libraries such as SciPy package[13], numpy[14], pandas[15], Scikit-learn[16], iPython[17]. In addition, matplotlib[18] and plotly[19] were used for visualization, and "R"[20] for data cleansing and wrangling.

2.2.1 "Style Detector Array" system architecture

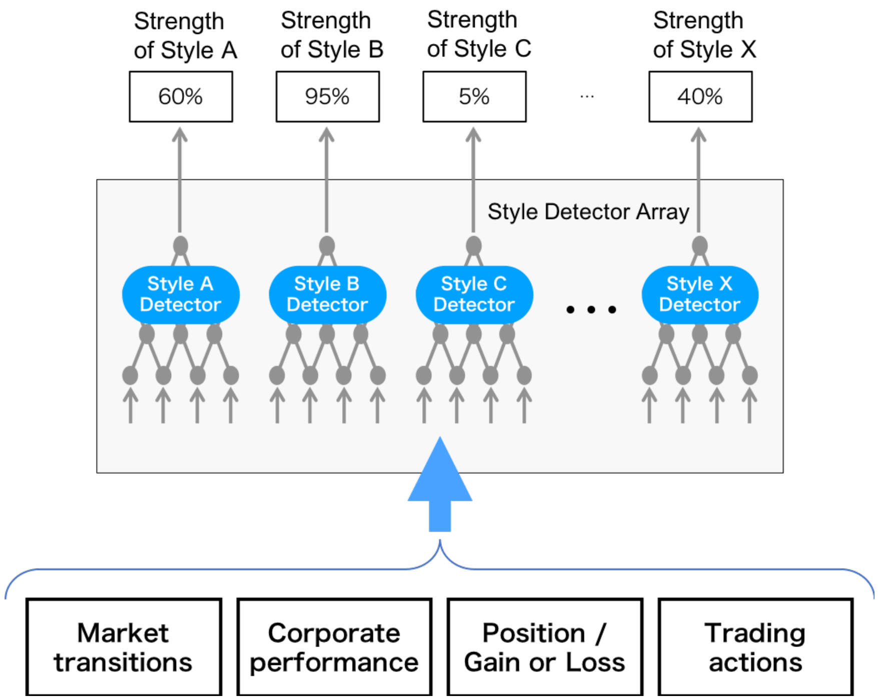


Figure 2.2: Style Detector Array

The Style Detector Array is composed of a number of detectors, which are individual identifiers that operate in parallel to evaluate the strengths of investment styles, which were determined as references. Each detector is implemented as a deep learning neural network (Figure 2.2). In addition to external scenarios such as the market environment and corporate performance trends, descriptions of the stocks that constitute the fund (market capitalization at each point of time and unrealized loss or gain) and daily trading actions are input as time-series, and an output is a vector representing the style of the fund manager. The vector is composed of values from each style detector, each representing the degree of similarity of the respective typical investment styles.

The generation of training data with virtual fund managers

To give the Style Detector Array the ability to classify investment behaviors into reference styles prepared in advance, the neural networks that constitute the detectors must be trained with investment behavior data paired with the reference styles. However, investment behavior data that "purely" reflects these reference styles does not exist in reality. Therefore, virtual trading data was generated with a simulation, using virtual fund managers implemented with investment style logic. The virtual trading data obtained in that way was used to train the Style Detector Array.

For this prototype, the eight investment logic elements listed below were set as strategies for the reference styles in the Style Detector Array. In setting the elements, consideration was given to variety, and what is believed to be typical in asset management, including elements that make decisions based on stock price, corporate performance and other factors that are considered to be fundamental, decisions that include window size determined by past time series trends rather than simple numerical values at an arbitrary point in time, those that involve only a particular small number of stocks, and those that involve stocks that extend across analysis universes.

- **High Dividend**
  A strategy of selecting and holding stocks that yield high dividends.
- **Minimum Volatility**
  A strategy of selecting and holding stocks that have exhibited low volatility in the past 20 business days.
- **Momentum**
  A strategy of selecting and holding stocks that have had high price increases over the past 20 business days.
- **Value**
  A strategy of selecting and holding stocks that have a low PBR.
- **Growth**
  A strategy of selecting and holding stocks that have a high PER.
- **Quality**
  A strategy of selecting and holding stocks that have a high operating cash flow to capitalization ratio.
- **Fixed Weight**
  A strategy of setting a target portfolio based on the equally weighted market capitalization of all stocks.
- **Technical**
  A strategy of swapping stocks according to patterns of change in long-term and short-term moving averages.

Virtual trading data was generated by simulating the trading behaviors of virtual managers that each adopt one of the trading strategies described above, based on historical data on market trends and corporate performance trends for the time period from November 1, 2005 to August 3, 2017.

Training style detectors with training data

Next, the detectors that constitute the Style Detector Array were trained using the virtual trading data sets from the respective virtual fund managers. The input to the neural networks of the detectors were factors related to the market and corporate performance (stock price, volume, dividends, PBR, PER, capital investment, free cash flow, etc) and factors related to the fund (position, unrealized loss or gain for each stock and trading actions, etc), provided for all of the stocks that constitute the universe, with a window of a number of days extending into the past. The assumed universe here consists of 100 stocks. There are 19 factors associated with each stock, and the window is set to 40 days, so the neural network input layer consists of 76,000 units (a rank-3 tensor of $40 \times 19 \times 100$). Of the virtual training data for each virtual manager, the data for the period from November 1, 2005 to February 24, 2015 is used in order to train the Style Detector Array, so that only the respective detector responds. The rest of the data for the period from

February 25, 2015 to August 3, 2017 was reserved as a validation dataset for evaluating the generalization ability of the trained detectors.

Incidentally, it is generally known that when neural networks are trained in complex problems, the output is sensitive to the values of parameters that are initially set at random at the beginning of training[21, 22, 23]. For that reason, the ensemble average of the outputs from 20 individually trained neural networks was taken as the output of the Style Detector Array. From the output of the trained Detector Array when the respective virtual trading data was applied as an input (Figure 2.3), we can see that the system can be trained to properly perform a rough discrimination of the eight reference styles, including for the time period of the validation data that was not used in the training.

(a) High Dividend (b) Minimum Volatility (c) Momentum

(d) Value (e) Growth (f) Quality

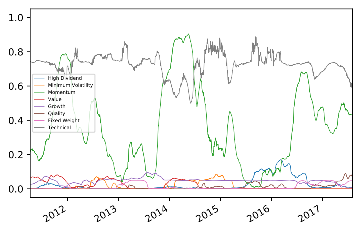
(g) Fixed Weight (h) Technical

Figure 2.3: Detector outputs after training

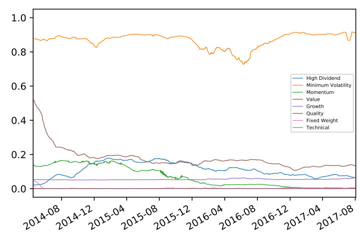
2.2.2 An analysis of domestic equity funds using the Style Detector Array

After the Style Detector Array was trained with the virtual trading data generated by the virtual fund managers to sufficiently discriminate the reference styles, data from real, existing domestic equity funds was used to evaluate the Style Detector Array. From the results of 10 domestic equity funds for which sufficient data was available, distinct "styles", varied from fund to fund, were evident, and temporal changes of style even within a single fund was observed (Figure 2.4).

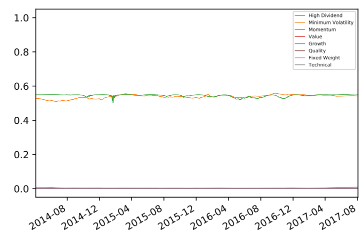
Below, the analysis for *fund"V"* and *fund"D"* are used as examples to dig deeper into the results.



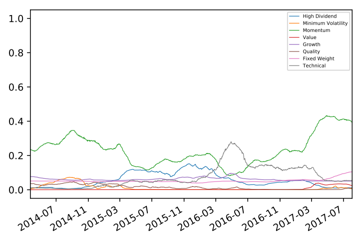
(a) *fund"J"*



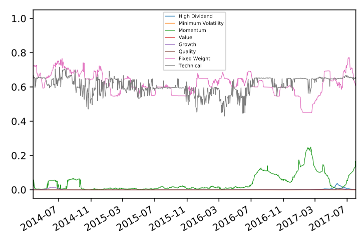
(b) *fund"B"*



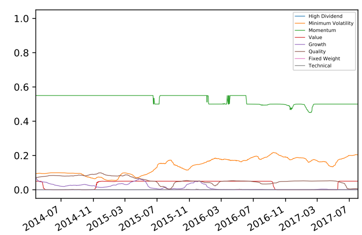
(c) *fund"O"*



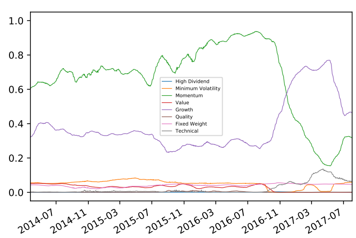
(d) *fund"Q"*



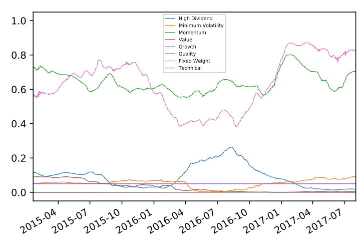
(e) *fund"L"*



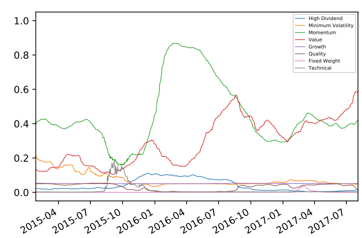
(f) *fund"E"*



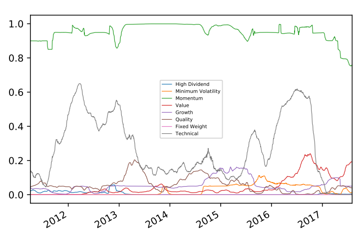
(g) *fund"V"*



(h) *fund"D"*

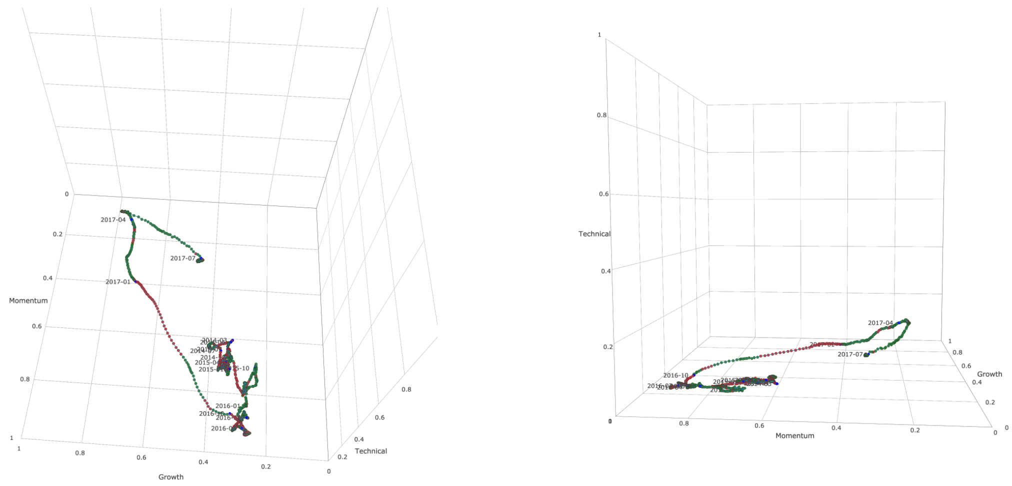


(i) *fund"X"*



(j) *fund"N"*

Figure 2.4: An analysis of a domestic stock funds using the Style Detector Array



(a) (b)

Figure 2.5: The trajectory in manager behavior phase space for *fund"V"*

Case Study 1: Analysis for fund"V"

The fund analyzed here mostly handles small growth stocks with distinct management policy and style.

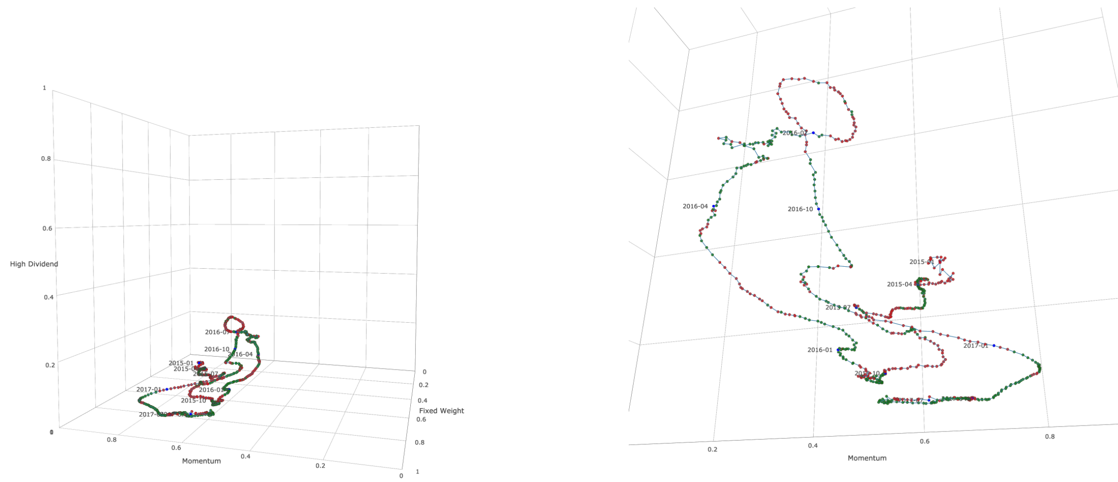
With a closer look at the time axis in Figure 2.4g, we first see that the attribute for the Momentum style is strongest, followed by the Growth style. The order of the styles reverses around November 2016, with Growth moving into the lead. At the same time, the Technical style appears as a strong component for the first time. The strengths of the other attributes remained low for the entire time period. The fact that the Growth style remained relatively strong throughout the time period is consistent with the fact that the fund deals mainly with small-capital growth stocks. These trends in the output of the Style Detector Array reveal a change in the investment style of the fund around November 2016. That fact was confirmed by the *fund"V"* management, who explained that the composition of the managers was changed around that time. The internal changes in the fund management were thus detected by the Style Detector Array as a style drift.

Mapping trajectories of vectors representing the "style" of the funds over the phase space is a useful visualization technique for gaining an intuitive understanding of temporal changes in investment style. Since the prototype Style Detector Array comprises eight detectors, the original phase space has eight dimensions. For visualization, however, the three most dominant output styles are displayed as a three-dimensional phase space. The trajectory for the investment styles of *fund"V"* in a three-dimensional phase space, with axes for Momentum, Growth, and Technical styles, is shown from two different perspectives in Figure 2.5. Triples that represent the output values of the Style Detector Array for the Momentum, Growth, and Technical style dimensions are plotted as coordinates in the phase space. The change in the strengths of Momentum and Growth that occurred around November is shown in the phase space as a movement of the trajectory from the lower right to the upper left in Figure 2.5a. This corresponds to the member changes within *fund"V"*. In Figure 2.4g, we see some fluctuation in the relationship between the Momentum and Growth styles in which the strength of Growth decreases and the strength of Momentum increases around April 2017. In the phase space representation in Figure 2.5, that behavior is observed as a bend in the trajectory that has a vertex at a point near April 2017.

Case Study 2: Analysis for fund"D"

This fund adopts a low-risk investment style in which investment styles are analyzed by a certain method according to market prices and the styles are nimbly rotated. This method emphasizes discretionary judgment by the manager that may go beyond pre-established strategies and policies.

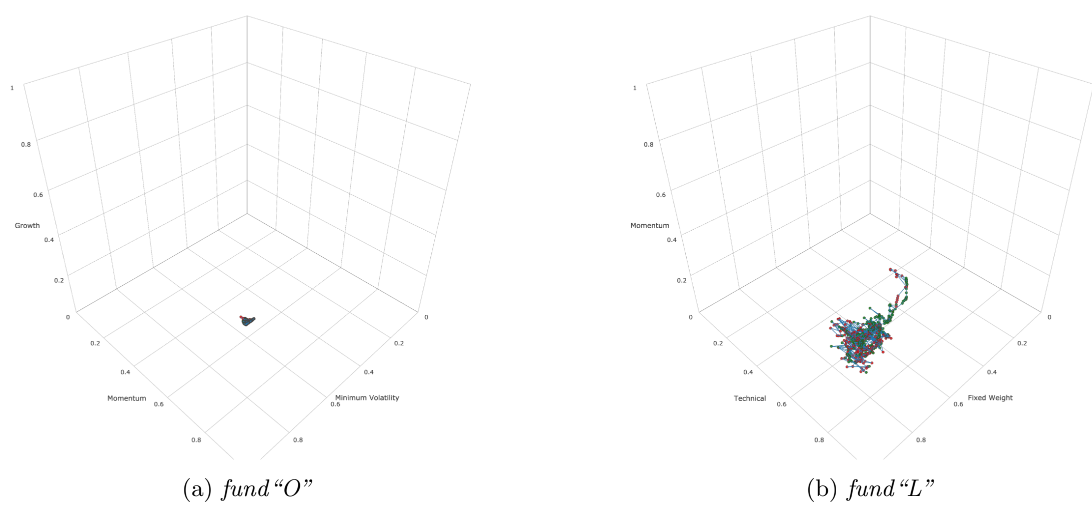
As we can see from Figure 2.4h, the Momentum, Fixed Weight, and High Dividend styles are generally dominant. However, major drifts in the phase space are observed in May 2015, from January to April 2016, and in November 2016, indicating changes in the strengths of the dominant styles.



(a) (b)

Figure 2.6: The trajectory in manager behavior phase space for *fund"D"*

The changes in investment style trajectories, in a phase space defined with the three dominant investment styles described above, are shown in Figure 2.6 from two perspectives. The rotating strategy characteristic of this fund is reflected exactly in the fact that particular combinations of strategy strengths do not remain stationary in the phase space in the long-term, and the trajectory tends to wander around. In contrast to the behavior of *fund"D"*, the trajectories in the phase spaces for *fund"O"* (Figure 2.7a) and *fund"L"* (Figure 2.7b) exhibit little to no change.



(a) *fund"O"* (b) *fund"L"*

Figure 2.7: The trajectories of funds with little change in investment style

2.3 Simultaneous visualization of multiple fund behaviors

Risk management is one of the most important aspects of GPIF. GPIF's commission of diverse fund managers, in anticipation of a diversity in investment strategy, leads to genuine robustness against unexpected market changes. Simultaneous visualization of multiple funds on the phase space is an intuitive way to recognize diversity and convergence of investment practices. Investment behaviors of all of funds analyzed in this study were represented as trajectories within a single phase space by using the t-SNE dimension reduction method[24].

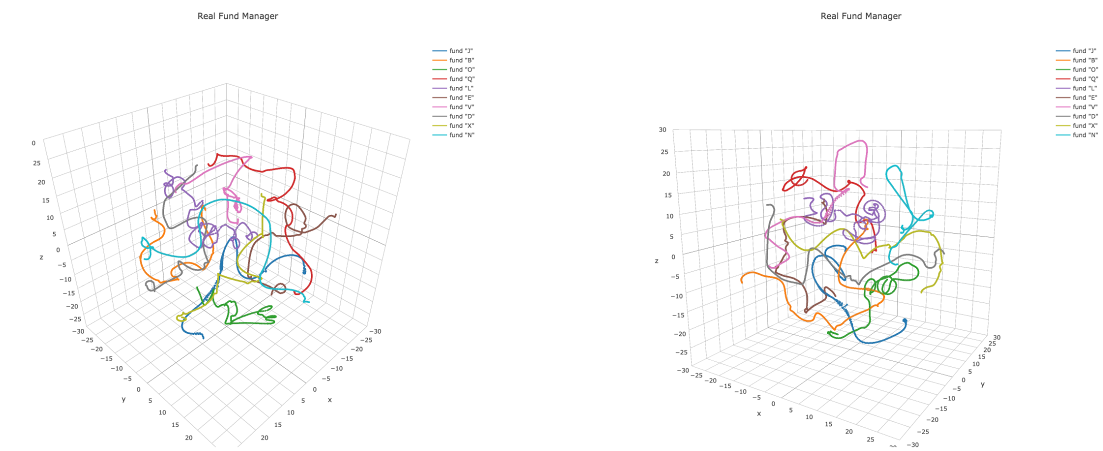


Figure 2.8: Actual fund management behavior: simultaneous displays with dimension reduction (2015/01 to 2017/04)

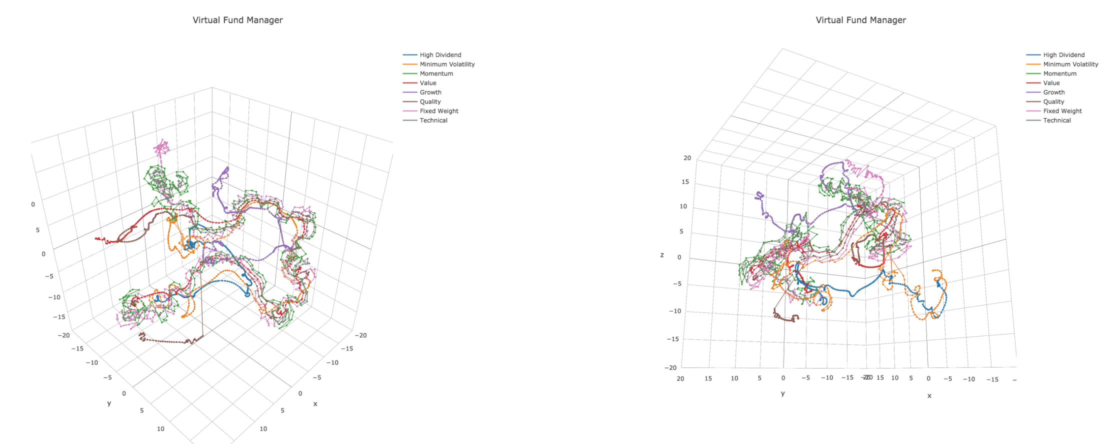


Figure 2.9: Virtual fund management behavior: simultaneous displays with dimension reduction (2006/06 to 2017/04)

Behaviors of actual domestic equity funds are represented as trajectories in a phase space (Figure 2.8). For this proof-of-concept prototype, only short-term data on actual domestic equity funds was used (from January 2015 to April 2017). To see changes in the robustness of the manager structure with respect to changing economic conditions, longer-term data extending over time periods that include important events such as the Global Financial Crisis, triggered by the subprime mortgage crisis in 2008, are required.

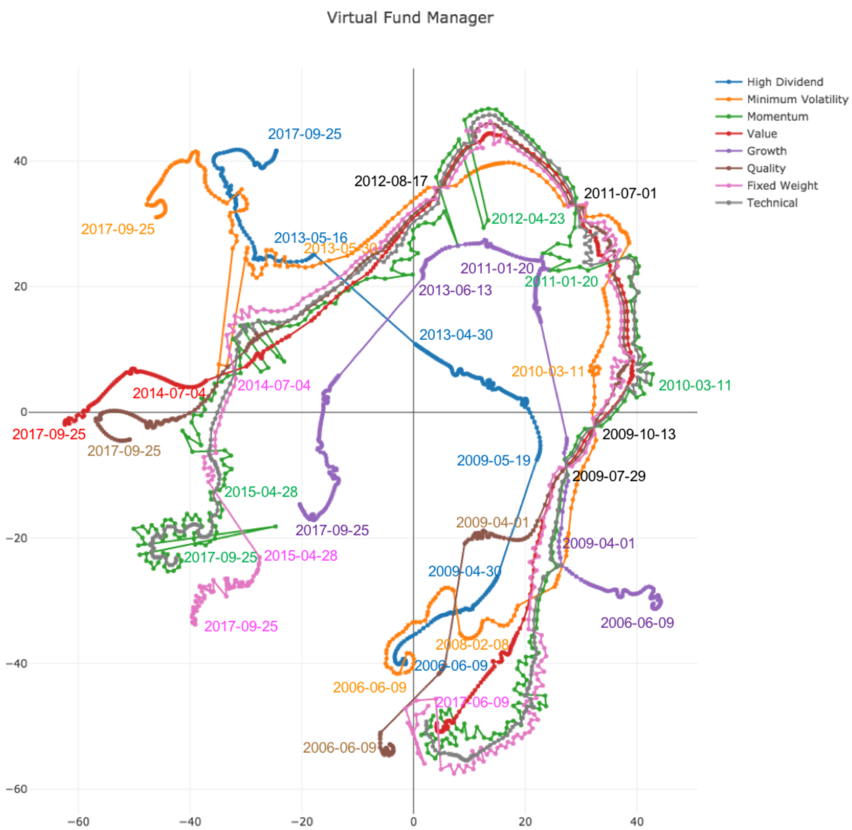
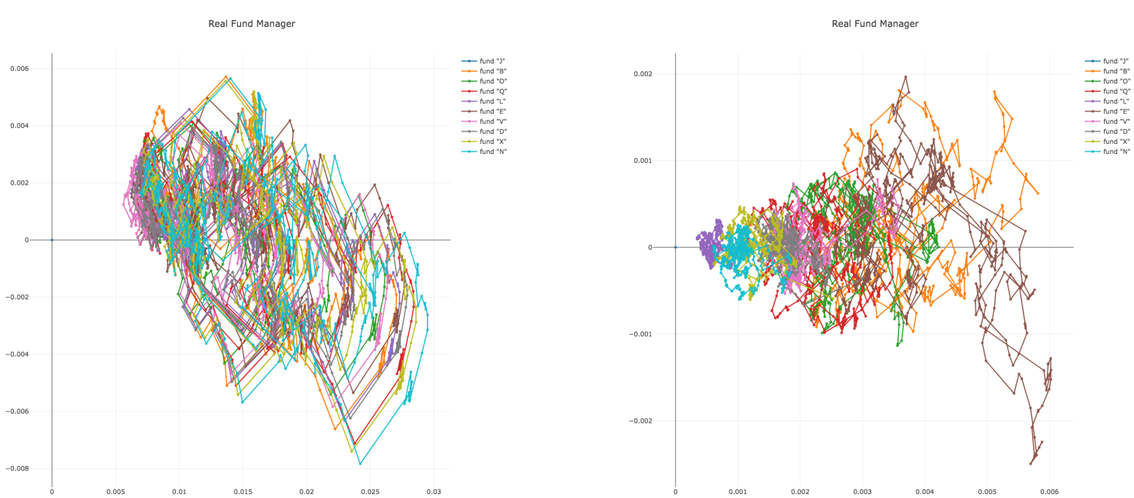


Figure 2.10: Virtual fund management behavior: reduced to two dimensions

A series of analysis using virtual trading data for the period from June 2006 to April 2017, which was generated by virtual fund managers as an alternative to the analysis of actual fund data, was performed to identify the possible detection of convergence. The behavior of the virtual funds represented as trajectories in a three-dimensional phase space is shown in Figure 2.9, and a representation of the same method in a phase space reduced to two dimensions is shown in Figure 2.10. In Figure 2.8, which was obtained using data from a limited time period, no notable change in structure is seen, but with the extended data, the trajectory is seen to converge and diverge over time. For example, the trajectories for the six virtual funds other than High Dividend and Growth begin to converge in April 2009 and nearly overlap until the Minimum Volatility trajectory separates at around May 30, 2013. Even after that, the trajectories for the other five virtual funds continue to be similar. Three virtual funds, Momentum, Fixed Weight, and Technical, are nearly the same over nearly the entire period, with an exception after April 28, 2015. We can see that the trajectories for Momentum and Technical, in particular, overlap until the end, even after April 28, 2015, and that the trajectories for High Dividend and Minimum Volatility mostly overlap, although there are periods of divergence from around April 30, 2009 to around May 16, 2013. On the other hand, the Growth virtual fund alone follows an independent trajectory that does not overlap almost at all in any of the time periods other than from April 1, 2009 to July 29, 2009.

The similarity of trajectories and phase space means that even the virtual fund managers that are designed to operate on the basis of different indicators and logic tend to behave in similar patterns according to background economic circumstances. This means that diversity of investment style is lost. These results suggest that if maintaining diversity is important, it is not sufficient to simply have fund managers that apply different investment styles. The fact that there is convergence of style with loss of diversity, even for virtual fund managers that blindly execute simply designed logic, leads us to expect that the same trend will be even stronger in the case of actual fund managers, where human factors such as optimism, fear and other emotions and arbitrary factors affect decisions.

2.4 Risk-return analysis

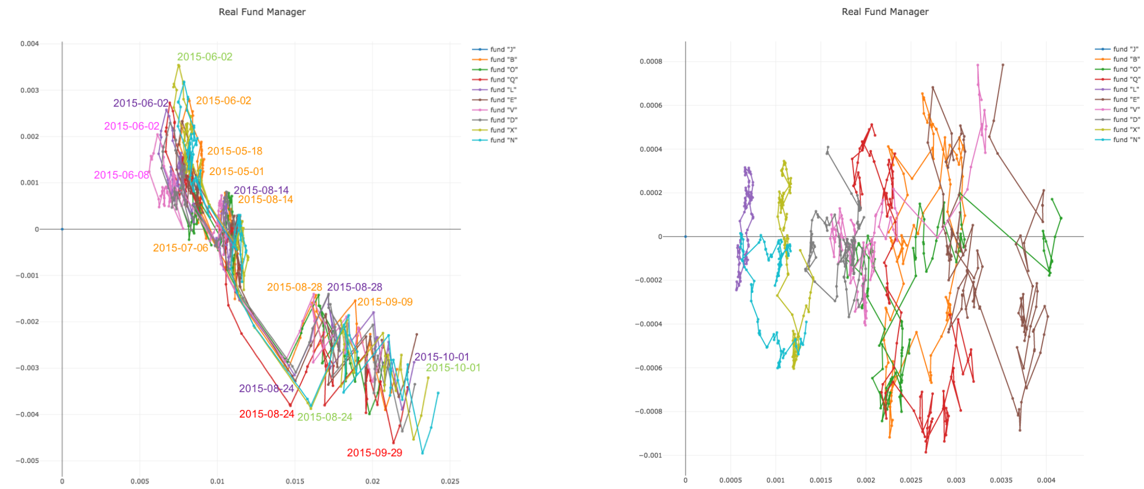


(a) Absolute return and risk

(b) Return and risk relative to benchmarks

Figure 2.11: Risk-return diagrams (2015/01/26 to 2017/04/06)

As a reference for comparison, we performed a traditional risk-return analysis on the domestic equity funds that we have been analyzing here. Taking the return to be the 20-day moving average of the daily return, and in the same way, taking the risk to be the standard deviation of the daily return for 20 days,



(a) Absolute return and risk

(b) Return and risk relative to benchmarks

Figure 2.12: Risk-return diagrams (2015/05/01 to 2015/10/01)

the relationship of the risk and return is represented as a graph, with the risk on the horizontal axis and the return on the vertical axis (Figures 2.11 and 2.12, where plots (a) show the relationship for the absolute return and risk and plots (b) show the relationship for the return, relative to a benchmark and risk).

Figure 2.11 covers all of the periods for which data is available from January 26, 2015 to April 6, 2017. Figure 2.12 covers only the period from May 1, 2015 to October 1, 2015, a period in which there was a large market movement. That period was selected because it includes the "China shock" of June and August 2015. In that period, the Shanghai composite index rose and peeked out after the Global Financial Crisis on June 12, 2015, and then fell by 30% over the following three weeks. On August 11, 2015, the People's Bank of China implemented the first major devaluation of the RMB in 20 years, triggering turmoil not only in the foreign exchange market, but in stock markets around the world. Then, the Shanghai stocks fell sharply again on August 18, in a worldwide chain of stock price weakness that lasted until August 26.

The overall impression from examining the risk-return diagrams is that the differences among the individual funds are not revealed clearly by the absolute return data (the graph on the left in each figure). In Figure 2.12a, which covers only the time period in which there were particularly large movements in the market, we can see clearly that the fund movements are synchronized. When looking at the data for returns relative to a benchmark (on the right side of each figure), however, some differences are evident. While *fund"K"* and *fund"M"* adopted a trading style that reduces risk relative to the benchmark, *fund"D"* and *fund"A"* adopted a style that permits risk levels that extend into a relatively high range. When the data is viewed relative to benchmarks in this way, we can see that there is some diversity in the fund investment styles.

2.5 Benefits provided by the Style Detector Array

The essential benefits provided by the Style Detector Array system are summarized below.

Direct analysis based on management behavior

Conventional tools, such as the Barra model or Aladdin, evaluate investment styles by examining changes in return, in terms of multiple factors and the sensitivity to each factor. On the other hand, our system directly analyzes the funds' behavior itself, which makes it possible to detect style drift earlier and more directly. Consistency checks between declarative descriptions and actual behavior, which have been done qualitatively, relying only on interviews so far, can be now conducted in an evidence-based manner. For GPIF, that can lead to significant improvement of the quality of interaction between GPIF and fund managers.

Flexibility of analysis

The trading styles that are set to serve as references for training the Style Detector Array are not limited to the eight used in developing the proof-of-concept prototype. The number of reference styles can be increased or decreased. Even entirely different reference styles can be used. Data on actual fund manager behavior can also be used for the references instead of the output of virtual fund managers that have simple trading styles. As with conventional tools, it is not necessary to be bound by measures set in advance. The analysis can be customized flexibly according to the purpose.

Separation of training phase and execution phase for data security

The training phase requires large amounts of computing resources and the use of cloud services and other external resources is assumed. Training the Style Detector Array, however, does not necessarily involve the use of highly confidential data, so information security concerns can be avoided. After training, on the other hand, the use of the Style Detector Array for analysis requires only the types of desktop or notebook computer that are readily available to individual users on the site. Thus, this analysis system can be operated with consideration given to both dependence on advanced computing resources (external cloud services, etc.) and data security.

Evaluation of AI-based funds

New kinds of funds utilizing the latest AI technology are emerging. Some AI-based funds may use deep learning that is largely a black box, which means the funds themselves may not be able to explain the behavior of the system in an understandable manner for humans. The Style Detector Array can be applied also to such kind of funds, because it can determine specific mixes of styles that are delivered from deep learning-based fund management. Therefore, the Style Detector Array might be well-adapted for analysis of these new funds, rather than the traditional human-based ones.

Chapter 3

Effects on the Business Models of Investment Companies

3.1 The impact of AI-based asset management systems on the business model of asset management companies

There are a range of AI applications on asset management. Here, we focus on a typical approach in which market data, as well as corporate performance and financial data, serves as the foundation for an AI-based asset management system. When introducing such a system, machine learning is applied to market data to predict stock prices and the direction and degree of change in stock prices. The alternative approach may be to detect the deviation of investment assets of interest from their expected prices, automatically in real time, which cannot be done manually.

What is important when using an AI system is that comprehensive prediction and detection will be possible for a wide range of assets, if not for an entire asset universe. Let's assume the case that such systems is introduced into the market on a large-scale. In this case, price deviations anywhere in the asset universe will be detected by multiple systems immediately, leading to the possibility of the deviations being eliminated in a short time. At the same time, the characteristics of the AI systems are determined by the data on which they were trained, so independently developed AI systems do not necessarily detect or predict price deviations in exactly the same way. As a result, there may be cases that the behavior of some of the AI systems will create such deviations. It is therefore difficult to determine whether temporarily detected deviations have natural causes or result from the effects of other AI systems.

AI systems continue to learn from data reflecting market characteristics that may result in the convergence of strategies over time. In this case, there is a possibility that a Nash equilibrium will be attained in an N-player game via the market. That means there is a high possibility that AI trading systems, which are based on deep learning or other statistical machine learning techniques and rely on market data or management indexes, will asymptotically approach index trading behavior.

Of course, not all AI asset management systems take the same approach, and there are methods of quickly predicting corporate performance from a wide range of data that extends beyond market information to include satellite images, etc., as seen in Two Sigma. Those must be treated as alternatives with a different framework. Nevertheless, the assumption is that most AI systems will be constructed on the basis of market data, corporate data, and information that relates to it, and in that case the scenario described here is probable. The business model will thus be determined by the balance between AI system investment that generates excess returns, and the management fee structure. Some funds will fail to develop and maintain systems that efficiently generate excess returns, and will be forced to either leave the market, or move towards index trading.

3.2 The effects of introducing AI systems within GPIF

Next, we consider the effects of GPIF introducing our prototype AI system. First, GPIF itself seems to have made major changes in manager selection and monitoring. For manager selection, GPIF should be able to obtain detailed analysis of investment styles, based on data from fund managers submitted in advance. We believe that enables dialog in presentations to be more precise, and backed up by data, thus establishing a highly effective manager selection process. Furthermore, the ability to analyze the investment styles of fund managers that are already under contract makes it possible to select the required investment styles, with understanding of the styles of all contractors. In monitoring, it is possible to analyze the actual management conditions of each fund manager, which should make it possible to increase efficiency. Clarity of communication in dialog is promptly responsive to changes in style.

The fund managers will be aware of the fact that GPIF has the ability to analyze investment styles in real time. The style and drift of individual fund managers will also be detected, and that may encourage a certain discipline among the fund managers. In addition, GPIF will be capable of analyzing the behavior of multiple fund managers from a comprehensive perspective – a bird's-eye view of fund manager investment styles. In reality, achieving excess returns constantly over a period of years is difficult to achieve, and the same is true for AI asset management systems. These circumstances create a need for fund managers to provide appropriate explanation. When asset management companies recognize that GPIF has the ability to independently analyze their investment styles and intends to continue development of even more advanced technology, they will recognize that they cannot justify their results with only qualitative explanations. As a result, asset management companies will take action to improve the efficiency of their investment process by introducing more sophisticated technologies, including AI, to explain their behavior and be accountable for their investment practices. We believe this will accelerate the use of AI-assisted asset management. At the same time, this will eliminate the dependence on individual persons from management strategies, and promote optimization. The transparency of the associated costs will also become high, as a result of reduced dependence on individual persons whose fair value is ambiguous at best. This sequence of developments will further promote the science and technology of asset management.

GPIF has access to data on a wide range of asset management companies that makes it possible to construct a huge and highly accurate database, covering a variety of management strategies, and a sophisticated management analysis system using the data. First, such a system benefits GPIF by contributing to the construction of a more sophisticated manager selection system. Second, it benefits the asset management industry as a whole if GPIF's insights can be shared with fund managers to improve their performance and risk management.

Chapter 4

Future work

There are obvious next steps beyond the proof-of-concept work described in this report. A brief discussion on some of options is described here.

Extension to the full universe of analysis

In developing the prototype described here, we performed various types of analysis on just 100 stocks that cover a variety of industrial sectors, selected from among 3600 or so stocks listed on the domestic equity market. We decided to limit the universe to a small size in order to ensure a large number of trial and error runs could be carried out within a limited schedule, with a low computation time. The obvious next step is to construct a Style Detector Array with the scope expanded to the full universe of analysis, which includes all listed stocks. That would enable the inclusion of funds that handle mainly stocks of small and medium capitalization, which were not well represented in the analysis presented here.

Elaborating the virtual fund managers

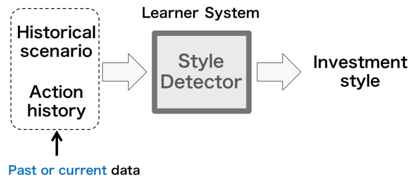
The logic of the eight investment styles that we defined for the virtual fund managers in the prototype were extremely simple and unrefined. The main purpose of this research was to verify operation of the framework and the results we obtained were sufficient for that purpose, but logic that is more precise and better fits the sense of the workplace is probably required for practical operation.

Expansion to different asset classes

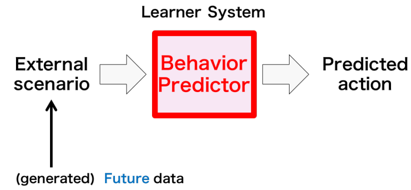
The proof-of-concept prototype described here was constructed on the basis of domestic equity funds because the data was easy to obtain and shape. However, this work can also be extended to other asset classes that comprise traditional assets other than domestic equities, including foreign equities, and foreign and domestic bonds. When doing so, it is necessary to take into accounts factors that are particular to each asset class. Considering foreign asset classes for example, the factors that must be considered in data shaping, in other preparatory work, and in implementing fund manager logic, include regional differences such as social stability, currency exchange rates, and so on. Bonds, the redemption term, interest rate, and issuer credibility (rating information) of bonds must be considered. Although some specific refinements related to the target are needed, the overall framework established by our prototype can be utilized with little modification when asset classes are extended.

From past analysis to future simulation

One ambitious direction is to try to model the fund managers' trading behavior, and to use the models for prediction of behavior under various scenarios. The current prototype has been designed and used as a learner for identifying investment styles based on past fund manager behavior, as shown in Figure 4.1a.



(a) Current prototype: Learner system as a style detector



(b) Further development: Learner system as a behavior predictor

Figure 4.1: From past analysis to future simulation

Going beyond such mere backward analysis of the past, we would like to move in the direction of using the learner as a part of the prediction system that predicts fund manager's behavior under external scenarios (Figure 4.1b).

Up to now, the evaluation of the risk/return ratio of a fund has basically been a matter of back-testing historical data, and stress testing also goes no further than the level of mechanically calculating the effects of a given scenario on the current portfolio. However, once we obtain a predictive model of fund manager's behavior, it becomes possible to use simulation for predictive evaluation of a fund manager's trading actions under arbitrary scenarios, as well as the prediction of performance and risk characteristics. Doing so should make it possible to carry out genuine forward-looking risk/return evaluation and stress testing, rather than relying only on past records, which we believe would lead to the construction of a more robust manager structure.

In pursuing this, we believe it is also useful to consider incorporating a methodology for generating virtual scenarios for external factors, such as market environments, which are a prerequisite for predictive simulations.

Chapter 5

Conclusion

This study focused on "manager structure" development and maintenance, a core activity of GPIF, as a target for the possible application of AI technologies. A proof-of-concept prototype Style Detector Array system was shown to be effective in detecting differences of style among fund managers, and how they can change over time. The Style Detector Array was developed using deep learning neural networks. The Style Detector Array was trained with data generated by virtual managers and then it was evaluated using data on the actual behavior of domestic equity funds. The results clarified the difference of investment styles among the funds, and that styles varied over time, even within a single fund. Some style drifts detected by the Style Detector Array were consistent with the timing of changes in manager composition in the management of actual funds.

In a simultaneous visualization of the behaviors of multiple funds, as trajectories in a single phase space using dimension reduction, the trajectories were seen to converge and diverge over time. The similarity of trajectories in phase space means that investment behaviors became similar under certain background economic circumstances, and the diversity of the manager structure was lost, even for virtual managers that were designed for different investment behaviors. This result suggests that, if maintenance of diversity is a top priority, it is not sufficient to compose a manager structure with funds that have different investment styles and that other means and measures are needed.

In the work reported here, we used a prototype Style Detector Array to directly analyze fund management behavior, enabling evidence-based, prompt analysis of investment styles. We believe the result is significant for GPIF because this enables it to acquire an evaluation and selection method that is more accurate, providing a path towards the composition of a more robust manager structure. In addition, this capability of GPIF may impact the strategy of asset management companies, upon their recognition of GPIF's capability to detect and monitor investment behaviors of fund managers more precisely and quantitatively.

In addition, the widespread adoption of AI by asset management companies may lead to a Nash equilibrium in an N-player game via the market. Under such a situation, AI asset management systems built on the foundation of deep learning and other forms of statistical learning will asymptotically degenerate to index investments. We suggested that, in such a scenario, business models would be determined by the balance between the cost for developing AI system and the management fee structure. Funds that fail to develop and maintain effective systems will either leave the market or trend towards index trading. Combined with the possible introduction of an analysis system based on the prototype that we produced in this research, GPIF will have the ability to independently analyze the investment style of various asset management companies, and will be able to monitor the status of alpha-generating activities, probably AI-based funds, and the tendencies of inter-fund convergence and divergence of investment behaviors. Asset management companies will soon notice the fact that GPIF is equipped with such capability, and intends to move forward with even more advanced technology. The result would be increased transparency in the explanation of performance and fee structure within asset management companies. That series of developments would promote the application of science and technology in asset management that would contribute to the stable and robust performance of GPIF.

Bibliography

[1] Government Pension Investment Fund (GPIF). Annual report fiscal year 2016, 2016.

[2] Government Pension Investment Fund (GPIF). Government Pension Investment Fund (GPIF), Summary of the 3rd Meeting of Board of Governors (in Japanese). http://www.gpif.go.jp/operation/management/pdf/keieiiinkai_306.pdf, November 2017.

[3] Government Pension Investment Fund (GPIF). Government Pension Investment Fund (GPIF), Summary of the 4th Meeting of Board of Governors (in Japanese). http://www.gpif.go.jp/operation/management/pdf/keieiiinkai_405.pdf, December 2017.

[4] David Weisberger and Paul Rosa. Automated equity trading: The evolution of market structure and its effect on volatility and liquidity. Technical report, Two Sigma Securities, 2013.

[5] Top page of Two Sigma. https://www.twosigma.com.

[6] Michael Markov. The law of large numbers: An analysis of the renaissance fund. Technical report, Markov Processes International, September 2007.

[7] Top page of Renaissance Technologies. https://www.rentec.com/.

[8] Samer Obeidat, Daniel Shapiro, Mathieu Lemay, Mary Kate MacPherson, and Miodrag Bolic. Adaptive portfolio asset allocation optimization with deep learning. *International Journal on Advances in Intelligent Systems*, Vol. 11, No. 1&2, pp. 25–34, 2018.

[9] Samer Obeidat. Five ways artificial intelligence is disrupting asset management. https://www.entrepreneur.com/article/312672, April 2015.

[10] Top page of Investifai. https://www.investifai.com.

[11] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.

[12] François Chollet, et al. Keras. https://keras.io, 2015.

[13] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001.

[14] Travis E. Oliphant. *Guide to NumPy*. Trelgol, 2006.

[15] Wes Mckinney. pandas: a foundational python library for data analysis and statistics. In *Python for High Performance and Scientific Computing*, January 2011.

[16] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. Scikit-learn: Machine learning in python. *J. Mach. Learn. Res.*, Vol. 12, pp. 2825–2830, November 2011.

[17] Fernando Pérez and Brian E. Granger. IPython: a system for interactive scientific computing. *Computing in Science and Engineering*, Vol. 9, No. 3, pp. 21–29, May 2007.

[18] John D. Hunter. Matplotlib: A 2d graphics environment. *Computing in Science and Engineering*, Vol. 9, No. 3, pp. 90–95, 2007.

[19] Plotly Technologies Inc. Collaborative data science. `https://plot.ly`, 2015.

[20] R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2014.

[21] Yi Shang and Benjamin W. Wah. Global optimization for neural network training. *IEEE Computer*, Vol. 29, No. 3, pp. 45–54, 1996.

[22] M. S. Iyer and R. R. Rhinehart. A method to determine the required number of neural-network training repetitions. *IEEE Trans. Neural Networks*, Vol. 10, No. 2, pp. 427–432, 1999.

[23] Akarachai Atakulreka and Daricha Sutivong. Avoiding local minima in feedforward neural networks by simultaneous learning. In Mehmet A. Orgun and John Thornton, editors, *Australian Conference on Artificial Intelligence*, Vol. 4830 of *Lecture Notes in Computer Science*, pp. 100–109. Springer, 2007.

[24] Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-SNE. *Journal of Machine Learning Research*, Vol. 9, pp. 2579–2605, 2008.

Reporter:
Takahiro Sasaki
Hiroo Koizumi
Takao Tajiri
Hiroaki Kitano

Sony Computer Science Laboratories, Inc.
Takanawa Muse Bldg. 3F,
3-14-13, Higashigotanda, Shinagawa-ku,
Tokyo, Japan 141-0022 Tel: 03-5448-4380