Investment Style Recognition

... reviewing research reports from financial institutions about recognising investment style of fund managers

- finding similarities in data about investment styles for predefined investment styles or not
- AI research by Japanese Pension Plan; GPIF(Government Pension Investment Fund)

GPIF

- The research was oriented toward harnessing the power of AI for one of GPIF's core functions, which is maintaining the manager structure.
- This function means to oversee the allocation of GPIF's pension assets among a diversified set of fund management firms (hereinafter called "funds" or "managers").
- The aim of the research was to provide more sensitive tools for detecting whether the investment behaviours of funds are kept consistent with their prospectus, and if not so evaluate whether this deviation is appropriate or need to be watched carefully.

Research Progress

phase	report	description
1st Phase (2017-2018)	summary report (ref.1)	- research title : "A Study on the Use of Artificial Intelligence within Government Pension Investment Fund's Investment Management Practices." - goal : detect investment style based on widely-used traditional factors - proof-of concept prototype system called SDA (Style Detector Array) based on deep learning was developed
2nd Phase (Oct.2018 - Oct. 2019)	interim report (ref.2)	- research title : "A Study on the Use of Artificial Intelligence for Learning Characteristics of Funds' Behavior" - goal : detect resemblance of fund manager's characteristics - developed a system called 'Resembler' as an extension and an application of SDA - also apply SOM (Self-Organising Maps) as a means of visualising fund characteristics

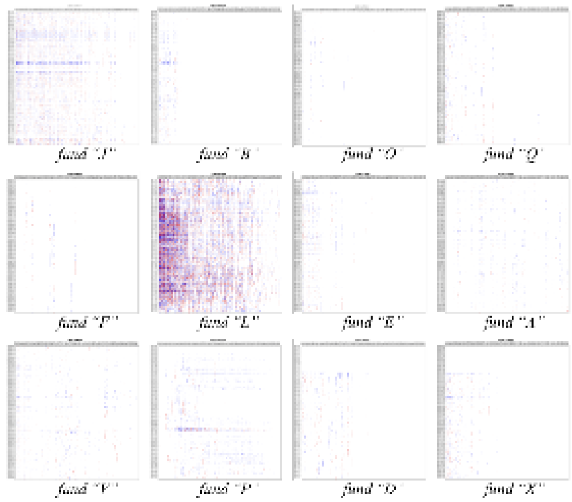
2nd Phase
 	
(Jan.2020 - Jun. 2020)
	final report (ref. 3)	- developed into two applications
		  - "Self-resemblance"
		    - detection of fund style change by providing past investment behaviour patterns
		  - "Mutual-resemblance"
		    - assessment of similarities among funds, primarily for fund selection

Model

SDA (Style Detector Array)

fund investment style

- the differences in investment style appear in the daily trading behaviours
  - i.e. fund deals mainly in small or large market capitalisation stocks (Fund 'P' and 'B')
  - quantitative style in which investment targets are bought and sold mechanically within a wide range based on mathematically determined decisions (Fund 'L')
  - discretionary method in which stocks are selected manually on the basis of careful investigation (Fund 'O')
  - intensive trading activity tells additional capital for investment or a large-scale re-allocation of assets

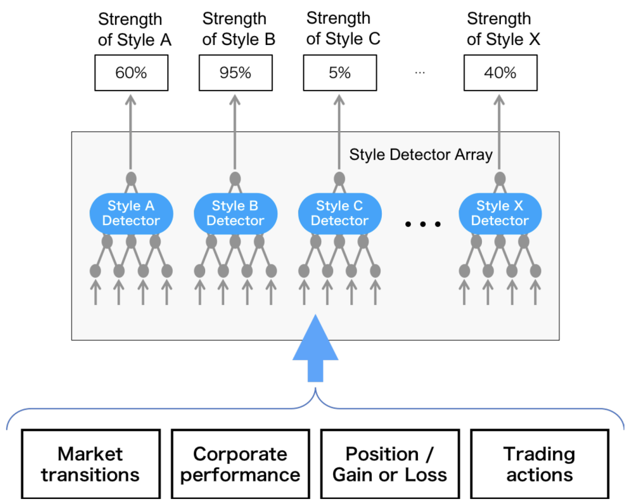


[A visualisation of fund manager trading behaviour]

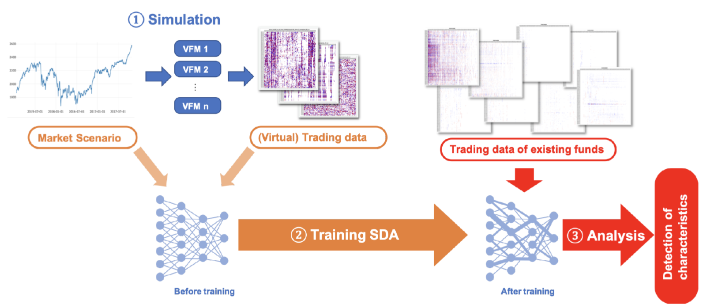
- x-axis : individual stocks in descending order of market capitalisations at the end of March 2018
- y-axis : days from April 2014, to March 2018
- blue dot indicates a buy trade and red dot indicate a sell trade

SDA (Style Detector Array)

- investment style analyser using deep learning technology
- composed of a number of detectors
- detector is 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
- input (time-series data)
  - external scenarios
    - market environment
    - corporate performance trends
    - descriptions of stocks that constitute the fund
      - such as market capitalisation at each point of time
      - unrealised loss or gain
  - daily trading actions
- output
  - a vector representing the style of the fund manager
  - is composed of values from each style detector, each representing the degree of similarity of the respective typical investment styles
- tools
  - Python
    - deep learning libraries : tensorflow, Keras
    - complementary libraries : SciPy, numpy, pandas, Scikit-learn, iPython, matplotlib and plotly
  - R for data cleansing and wrangling



[Style Detector Array]

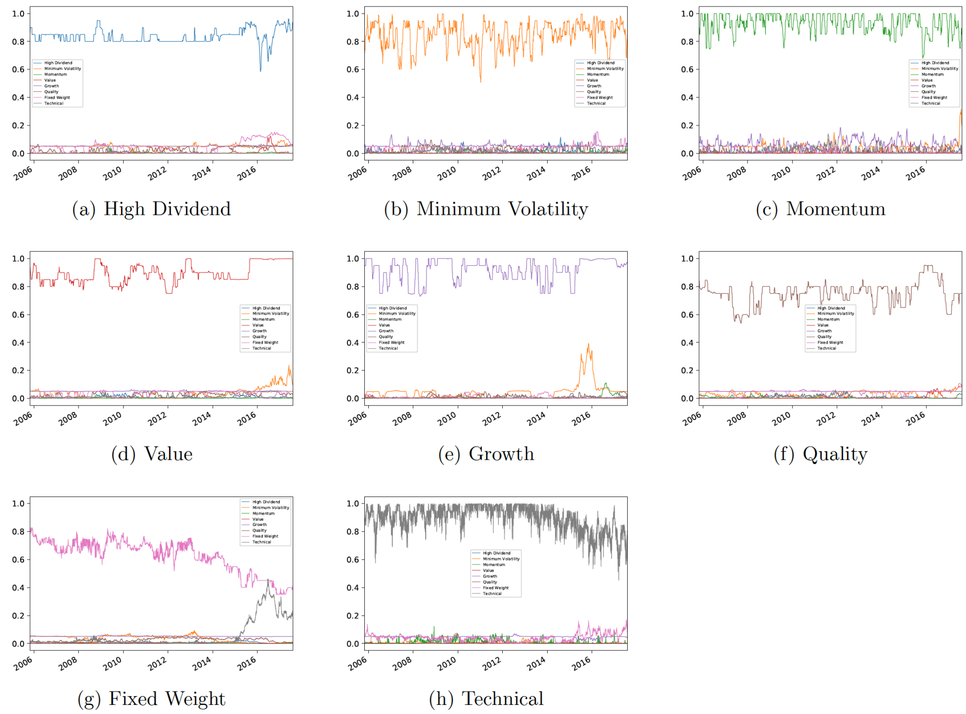


[Flow of analysis made by SDA]

1. generate virtual trading data through simulation of virtual fund managers
2. train SDA using virtual trading data
3. input actual trading data from real funds into the trained SDA for analysis

trading data for training

- need to trained with investment behaviour data paired with the reference styles
- generated virtual trading data with simulation using virtual fund manager
- virtual fund manager
  - implemented with investment style logic
  - based on a small universe of 100 Japanese equities selected mainly according to market cap
  - for the time period from November 1, 2005 to August 3, 2017
  - each adopt one of the trading strategies below based on historical data on market trends and corporate performance trends



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

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

(g) Fixed Weight (h) Technical

[Detector output after training]

investment logic elements (widely used traditional factors)

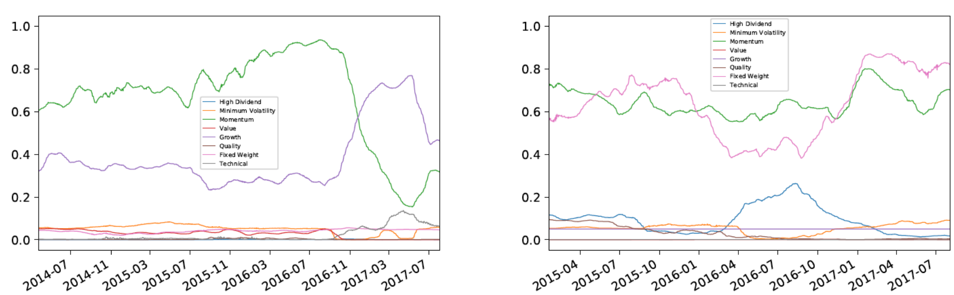
1. high dividend : a strategy of selecting and holding stocks that yield high dividends
2. minimum volatility : a strategy of selecting and holding stocks that have exhibited low volatility in the past 20 business days
3. momentum : a strategy of selecting and holding stocks that have had high price increases over the past 20 business days
4. value : a strategy of selecting and holding stocks that have a low PBR
5. growth : a strategy of selecting and holding stocks that have a high PER
6. quality : a strategy of selecting and holding stocks that have a high operating cash ow to capitalisation ratio
7. fixed weight : a strategy of setting a target portfolio based on the equally weighted market capitalisation of all stocks
8. technical : a strategy of swapping stocks according to patterns of change in long-term and short-term moving averages

Fund V

- mostly handles small-cap growth stocks
- attribute for the Momentum style is strongest followed by Growth style
- the Technical style appears as strong component
- the other attributers remained low for entire period of time
- strong growth : fund deals mainly with small capital growth stocks
- the order of the styles reverses around November 2016
- management who explained that the composition of the managers was changed around that time

Fund D

- mostly low-risk investment style
- the Momentum, Fixed Weight, and High Dividend styles are dominant
- major drifts are observed in May 2015, from January to April 2016, and in November 2016
- 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



(g) *fund "V"* (h) *fund "D"*

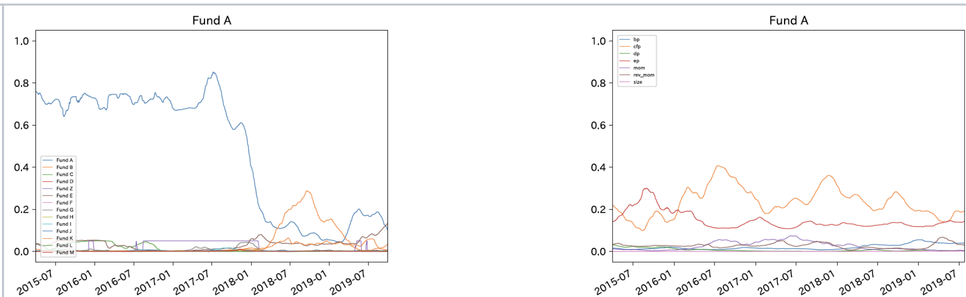
[An analysis of a domestic stock funds using the SDA]

- benefits
    - a blind test of style detection using actual trading behaviour data for 16 fund managers demonstrated that the system can properly detect the styles and drifts of each fund manager
    - style analysis could be performed in real time with an evidence-based approach, enabling GPIF to evaluate and select its fund managers more accurately
    - in addition, the system's visualisation capabilities were proven to be effective in identifying the spontaneous convergence of trading behaviours where most funding managers happen to trade similar items

- problem of qualitative evaluation in fund management firms (A study on the use of artificial intelligence within government pension investment fund)
    1. difficulties in avoiding arbitrariness and subjectiveness, and
    2. the dependence on a small pool of staff with specialised expertise

Resembler

- extension of SDA
- both Resembler and SDA have same underlying system architecture
- differing only in what data is used for training
- upgrade SDA
    - scaling up the universe of equities (100 to 1000 equities)
    - round-robin sampling for defining the universe
    - investment strategies adopted as reference styles
    - VFM as aggregated clusters of sub-VFMs
    - Smoothing by category during equity selection
- trained on actual trading data from actually existing fund
- output
    - a time-series of vector that track the resemblance of a fund to each of the actual funds used as references
    - i.e. its resemblance to Fund A



(a) Resembler (b) SDA

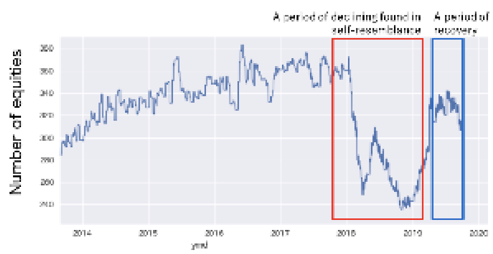


Figure 10. Case study 1. Evolution of number of equities held

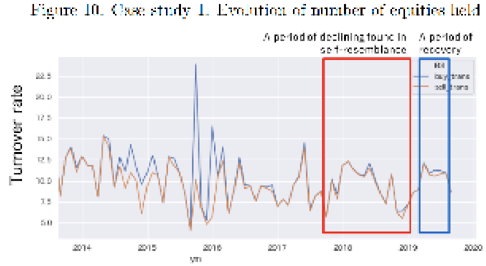
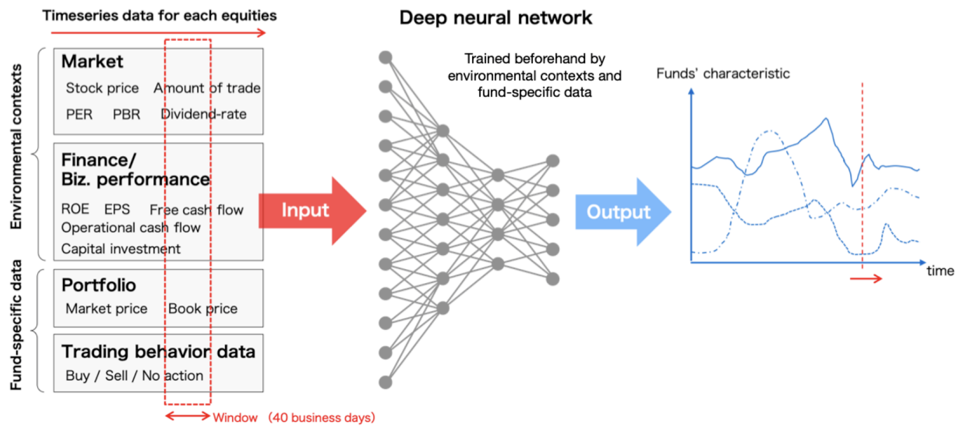


Figure 11: Case study 1: Evolution of turnover

<table>
<tr>
<td>
self-resemblance

- by comparing a fund to its own past characteristics and uniqueness
  - by including past investment behaviour patterns in the training data set
  - assess the degree to which a fund is maintaining consistency with its own characteristics and uniqueness
  - indicator of whether the fund has been changing its style
</td>
<td>



[Overview of Self-resemblance]
</td>
</tr>
<tr>
<td>
mutual-resemblance

- which is the degree of inter-similarity among multiple funds
  - trained with investment activity data of a certain number of funds selected as "anchors"
  - given the target fund of analysis, as an input,
  - outputs n-dimensional vector, whose attribute values express the degree of resemblance to the corresponding anchor funds
  - measuring the distance between these vectors for multiple funds, we can evaluate the mutual-resemblance among these funds
</td>
<td>

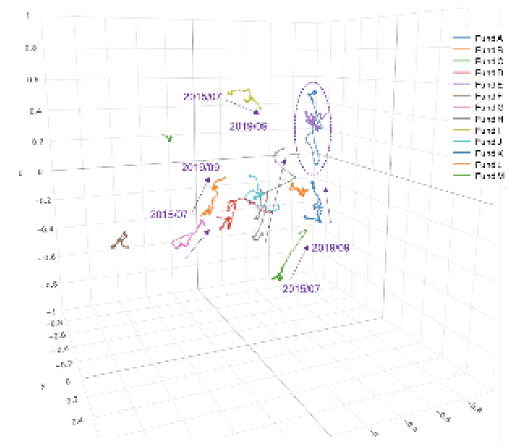


Figure 19: Changes in mutual resemblance for Japanese funds over time

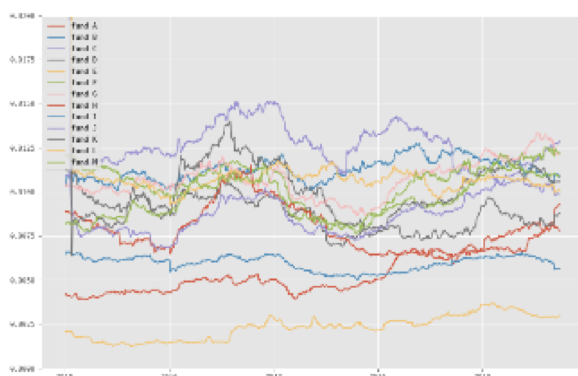
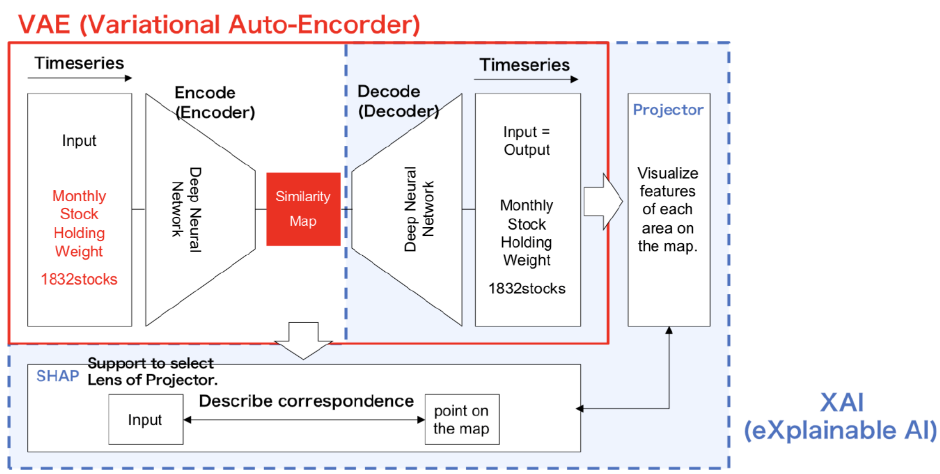


Figure 20: Changes in tracking errors
</td>
</tr>
</table>

[Overview of Mutual-resemblance model]

- encoder takes the holding weight in each equity as input
- and maps this as two-dimensional data
- map represent the similarity among funds
- projector
  - has a function that visualise the output to show features
  - such as equities, industries and risk factors on the map
  - serves as a lens to explore the map from various points of view
- SHARP (Shapley Additive Explanation value)
  - provides a significance indicator to guide the choice of lens
- VAE (Variational Autoencoder)
- XAI (Explainable AI) called SHARP (Shapley Additive Explanation value)



- benefits
  - detected potential style drifts that were not detected by Style Detector Array (SDA)
  - mutual-resemblance uncovers similarities among funds and their temporal changes assisting GPIF to maintain diversity of investment style

SDA vs Resembler

	SDA	Resembler
Universe size and asset class	Small universe (100 Japanese equities)	Large universe (1,000 Japanese or foreign equities)
Indices to be detected	Investment style based on widely-used traditional factors	Resemblance among fund managers' characteristics
Target of the development	Proof-of-concept prototype to validate the possibility of detecting investment style	New model and various system upgrades to support experimental implementations
Training data	Virtual trading data generated by VFMs	Actual trading data from existing funds
Target of detection	Objective indicators of “investment style” defined according to widely-used factor exposure	Relative indicators such as “resemblance to Fund A”
Assess fund behaviour	the third person perspective ▪ using commonly used investment styles ▪ actual investment style can be changed without affecting composition of style index which captured by SDA	the internal party perspective ▪ quantifying characteristics such as the uniqueness and consistency of how the fund is managed

Other improvement

Using ABCI for improvement

- training and execution of SDA and Resembler requires massive computing resources
- deploy system on a large-scale GPGPU cluster machine owned by the Japanese Government
- ABCI (AI Bridging Cloud Infrastructure) is provided by the National Institute of Advanced Industrial Science and Technology
  - ABCI is a computing cluster constructed from 1,088 nodes with GPUs
  - and provides an optimal infrastructure for parallel process execution of deep learning algorithms
  - iteration cycles which took a week or more to run previously can now be completed in about a day and a half

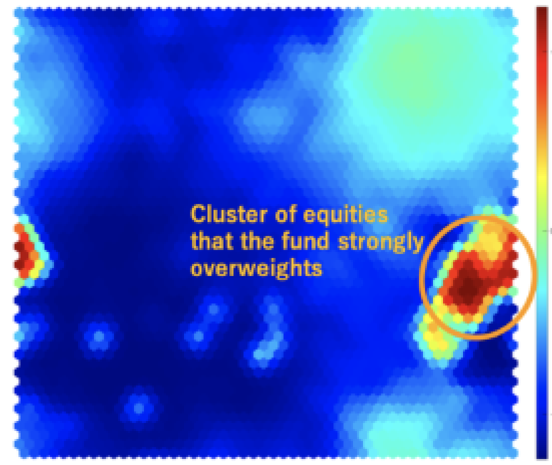
Self-Organising Maps

- method of creating intuitive visualisation of fund characteristics and changes
- type of neural network that maps a large number of individual items onto a 2D or 3D spaces according to the resemblance among those attributes
  - each of individual items has a certain number of attributes
- performs unsupervised clustering on the attribute vectors of complex high-dimensional data,
- and creates visual representations enabling humans to grasp trends and correlations that would be difficult or impossible to spot in the raw data

The basic flow of SOM is as follows:

1. An item is selected randomly from the dataset on which clustering is to be performed.
2. Compare the attribute vector of that item with the reference vector of each unit on the map and select the best matching unit.
3. Place the item in the best matching unit and adjust the reference vectors of the surrounding
   units to be closer to the item's attribute data vector.
4. Return to step 1.

then, the map will ultimately consist of clusters of items whose attribute vectors are most similar to each other



[Visualisation of holding patterns through SOM]

- the equities held by a certain fund are clustered
- red zones indicate equities which the fund overweight
- zooming in on the units, can see which specific equities are in the cluster and investigate the attributes corresponding to the units

Case Study

type	case study	application
fund selection	What is the key differentiator between existing Fund A and candidate Fund B? (Japanese equity funds)	mutual-resemblance
	Detection of change in portfolio manager at candidate Fund C (Japanese equity fund)	self-resemblance
monitoring of existing funds	Why did existing Fund D exhibit change in its investment behaviour during the 'Coronavirus shock' equity market?(foreign equity fund)	

Ref.

1. A Study on the Use of Artificial Intelligence within Government Pension Investment Fund's Investment Management Practices (Summary Report) 2018
2. An Interim Report on "A Study on the Use of Artificial Intelligence for Learning Characteristics of Funds' Behavior" 2019
3. A Study on the Use of Artificial Intelligence for Learning Characteristics of Funds' Behavior (Summary Report) 2020
4. GPIF - pionline article 2019