

Perez, Katarzyna; Szczyt, Małgorzata

Article

Classification of open-end investment funds using artificial neural networks. the case of Polish equity funds

Central European Economic Journal (CEEJ)

Provided in Cooperation with:

Faculty of Economic Sciences, University of Warsaw

Suggested Citation: Perez, Katarzyna; Szczyt, Małgorzata (2021) : Classification of open-end investment funds using artificial neural networks. the case of Polish equity funds, Central European Economic Journal (CEEJ), ISSN 2543-6821, Sciendo, Warsaw, Vol. 8, Iss. 55, pp. 269-284, <https://doi.org/10.2478/ceej-2021-0020>

This Version is available at:

<https://hdl.handle.net/10419/324548>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

ISSN: 2543-6821 (online)

Journal homepage: <http://ceej.wne.uw.edu.pl>

Katarzyna Perez, Małgorzata Szczyt

Classification of Open-End Investment Funds Using Artificial Neural Networks. The Case of Polish Equity Funds

To cite this article

Perez, K., Szczyt, M. (2021). Classification of Open-End Investment Funds Using Artificial Neural Networks. The Case of Polish Equity Funds. Central European Economic Journal, 8(55), 269-284.

DOI: 10.2478/ceej-2021-0020

 To link to this article: <https://doi.org/10.2478/ceej-2021-0020>



Katarzyna Perez 

Poznań University of Economics and Business, Poland
corresponding author: katarzyna.perez@ue.poznan.pl

Małgorzata Szczyt 

Poznań University of Economics and Business, Poland

Classification of Open-End Investment Funds Using Artificial Neural Networks. The Case of Polish Equity Funds

Abstract

In this study we utilise artificial neural networks to classify equity investment funds according to two fundamental risk measures—standard deviation and beta ratio—and to investigate the fund characteristics essential to this classification. Based on a sample of 4,645 monthly observations on 37 equity funds from the largest fund families registered in Poland from December 1995 to March 2018, we allocated funds to one of the classes generated using Multilayer Perceptron (MLP) and Radial Basis Function (RBF). The results of the study confirm the legitimacy of using machine learning as a tool for classifying equity investment funds, though standard deviation turned out to be a better classifier than the beta ratio. In addition to the level of investment risk, the fund classification can be supported by the fund distribution channel, the fund name, age, and size, as well as the current economic situation. We find historical returns (apart from the last-month return) and the net cash flows of the fund to be insignificant for the fund classification.

Keywords

open-end investment fund classification | equity funds | artificial neural networks | emerging market

JEL Codes

G23, C38, C45

1. Introduction

Open-end investment funds are publicly offered and classified in practice by ranking agencies such as Morningstar. The classification of these funds consists in assigning a given fund to a group of funds having the same investment style, which reflects the investment policy declared in the fund's prospectus and determines the level of investment risk taken by the fund manager. The style is visible in the name of the fund, so that the potential investor can identify and select a fund with an acceptable risk level for investment. Thus, the classification of funds facilitates their distribution and sale.

The current literature shows that managers of open-end investment funds often perform so-called style drift (K. C. Brown et al., 2015; S. J. Brown & Goetzmann, 1997; Chua & Tam, 2020; DiBartolomeo & Witkowski, 1997; Sha, 2020; Wermers, 2012). This

means that they add financial instruments representing a different investment style to a fund's portfolio (for example, they may buy shares for a debt fund, or they may buy bonds into an equity fund) in order to increase this fund's profitability. The higher return is sought in order to increase the competition against peers and to attract new potential investors to such a fund. However, such action often causes the risk level of the fund to be different than declared (higher or lower), and the fund itself becomes misclassified (S. J. Brown & Goetzmann, 1997). Inaccurate fund classification produces the wrong signals and misleads investors, who become lost in the very rich offerings of these funds nowadays. This especially concerns individual investors, who are the dominant group of fund participants, characterised by different levels of risk tolerance and financial education that influence their investment decisions (Jiang et al., 2020; Müller & Weber, 2010). For this reason, there is a need for deep diagnosis of the methods of classifying

open-end investment funds, which are key in their sale. New technologies come with help, including machine learning, which can be used as an objective tool for classifying open-end investment funds. It can be offered through robo-advisory, which is cheaper and less error-prone than stationary advisory (Jung et al., 2019). Thus, it can provide tangible benefits for both the demand and supply sides of this market.

The purpose of this study is twofold. First, it utilises machine learning tools to confirm the classification of open-end mutual funds, which is based on fund styles, reflected in funds' names and designation of investment risk levels and prepared by rating agencies. Second, it verifies whether any other attributes, such as historical returns, age, size, cash flow or the channel of distribution and the current economic conditions, may influence this classification in addition to investment risk. In this article, we are going to answer these questions in the case of open-end mutual funds representing one style: equity. The justification for this choice is as follows. Equity funds are the dominant subject of research on open-ended mutual funds in general, and certainly research on investment style. This is because these funds have the greatest volatility and the highest management activity among open-end funds, and therefore, the greatest possible temptation for management to change the style. Moreover, equity funds constitute almost half of the population of open-end investment funds in the world (Investment Company Institute (2020)), Fig. 1.1). In this study, we focus on equity funds from Poland. Contrary to the mature and widely investigated fund market of the USA, this market is an emerging one and it still has a low share of stock market capitalisation and the net asset value in GDP (Investment Company Institute [(2020)], Figure 1.10). Nonetheless, its growth potential is very high, if only because of the PLN 1.3 trillion (EUR 0.3 trillion) in local household savings, which, due to the zero interest rate on bank deposits, has been transferred richly to open-end investment funds. Growing individual investors' attention to the growing emerging fund market that consists of a wide offer of open-end investment funds justifies the need to verify whether these funds are classified correctly.

We base our study on a sample taken from the Morningstar Direct database. The sample consists of 4,645 monthly observations of 37 equity funds from the largest fund families registered in Poland from December 1995 to March 2019. We allocated funds to one of the classes generated using Multilayer Perceptron (MLP) and Radial Basis Function (RBF),

used to solve classification problems. Risk measures—that is, standard deviation and beta coefficient—are our classifiers. The results of the study allow us to confirm the legitimacy of using machine learning as a tool for classifying open-ended equity investment funds. Artificial neural networks provided more applicable results in the case of standard deviation as a classifier than did beta ratio. In addition to the level of investment risk, the sigma classification is supported by the fund distribution channel, the fund name and the current economic situation, as well as the fund age and its size. We find historical returns (apart from the last-month return) and the net cash flows of the fund as insignificant for the fund classification.

To our knowledge, this study is the first in which neural networks are used to verify the correctness of the classification of open-end investment funds. The literature is dominated by two approaches to studying the style drift. The first, proposed by Sharpe (1992), is the return-based analysis, which considers the volatility of fund returns and shows the manager's actions regarding the investment risk. The second, described by Daniel et al. (1997), is the holding-based analysis, which examines the detailed composition of a fund portfolio in terms of its changes and volatility. Our study complements the results obtained using both methods by classifying mutual funds according to risk measures and, additionally, various fund characteristics that may be significant for this classification. We classify funds using artificial neural networks, one of the methods of machine learning which is an effective research tool implemented lately in the area of financial markets and asset pricing (Gandhmal & Kumar, 2019; Gu et al., 2020; Pandurang & Kumar, 2019). However, in the case of studies on open-end investment funds, machine learning is still a niche utilised so far only to examine the quality of an open-end fund price prediction and not to study the quality of fund classification. For example, Indro et al. (1999) extend the pioneer research of Chiang et al. (1996) by predicting 1-factor Jensen's alpha of funds with different styles using multilayer perceptrons (MLP) and the GRG2 nonlinear optimiser. Wang & Huang (2010) predict the Sharpe index by comparing the backpropagation neural network (BPN) to the fast adaptive neural network classifier (FANNC). Pan et al. (2019) investigate the possibility of predicting the net asset value (NAV) of equity funds by use of the functional link artificial neural network (FLANN), and Rout et al. (2020) add technical extensions to it. They confirm that machine learning techniques generate better open-end investment fund price

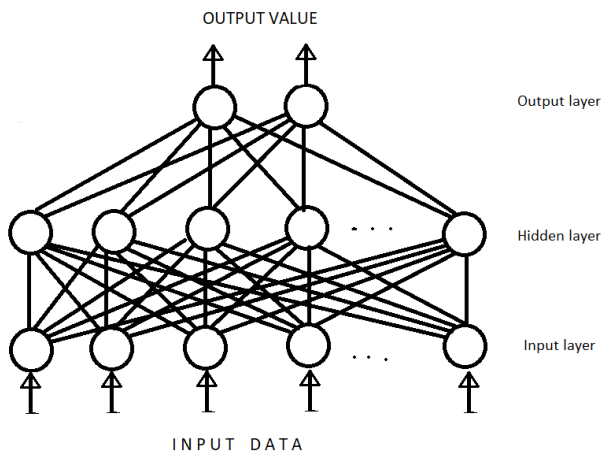


Figure 1. Diagram of an artificial neural network
Source: own elaboration

forecasting results than commonly used linear models. We hypothesise similar results, this time, though, as far as the classification of open-end investment funds is concerned.

Finally, in this study, the research gap is filled regarding the classification of investment funds from emerging European economies. There is much evidence of the style drift on the mature American market (Bams et al., 2015; S. J. Brown & Goetzmann, 1997; Cremers & Petajisto, 2009; M. Kim et al., 2000; T. H. Kim et al., 2005; Mason et al., 2012; Sensoy, 2009). Researchers provide some evidence of the style drift in emerging fund markets in China (Chua & Tam, 2020; Sha, 2020; Zhou et al., 2018). For India, Mohanti & Priyan (2018) show that the fund managers exhibit some level of active management and a good selection capability. However, the literature on the classification of European funds is fragmentary. Castellanos and Alonso (2004), and Moreno, Marco and Olmeda (2006) demonstrated manipulation of the risk of a portfolio of funds by a large proportion of investment fund managers registered in Spain. On the other hand, Zamojska-Adamczak (2005) showed similar behaviour for Poland, unlike Białous and Truszkowski (2009)'s research on Polish funds. They used Sharpe's (1992) methodology and indicated that the equity funds studied by them invested the accumulated capital in line with the declared investment policy, maintaining approximately 86% of shares in the portfolio. In this study we hypothesise a similar result: not shifting the style of equity funds, but carrying out an up-to-date investigation utilising a novel method to classify open-end investment funds. We also investigate which fund attributes are important to this classification. Our

findings support retail investors who might not be equipped with the tools to track the possible style drift, but who may deduce which funds deviate from the defined fund classification by analysing the funds' names, channels of distribution, age, size and returns, as well as the current economic conditions.

The remainder of the article is organised as follows: in section 2 we characterise artificial neural networks. In section 3 we present the data, variables and research procedure for classifying mutual funds using artificial neural networks. In section 4 we discuss the results and in section 5 we conclude.

2. Research methods

In order to classify mutual funds according to their investment style, the method of artificial neural networks (ANN) is applied. Artificial neural networks are information (data) processing systems. The main idea behind the operation of ANN is to reproduce the operation of the human nervous system, where signals are transmitted between neurons. Artificial neural networks are made up of layers of artificial neurons (see Figure 1). The input layer contains the number of neurons equal to the number of dependent (explanatory) variables used in the analysis.

Next, in the artificial neural network, there are hidden layers of neurons (there may be one, several or a dozen of these layers, depending on the complexity of the analysed problem). Neurons in the input layer are connected to the neurons in the first hidden layer by means of a weight system, and the neurons in the adjacent hidden layers are also connected. Thanks to weights, signals (values) coming from neurons are transformed before being transferred to other neurons. As each neuron obtains signals from many other neurons, these signals are aggregated in the neurons (see Figure 2). The most common types of aggregation are:

- linear aggregation, where the aggregate value is a linear combination of the inputs:

$$z = \sum w_i x_i$$

- radial aggregation, based on the distance of the vector of variables x from the vector of weights in:

$$z = \sum (x_i - w_i)^2$$

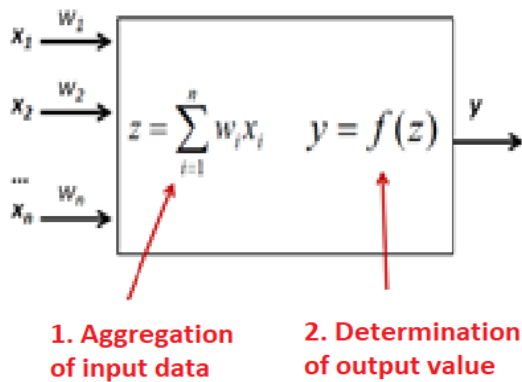


Figure 2. Neuron diagram
Source: own elaboration

As this aggregation is a measure of the similarity between weights and values of the variables, it is often used in classification problems.

The output value of the neuron is determined by an activation function, which can be any kind of function. S-shaped, linear, or Gaussian are the most commonly used.

The final layer is output, which provides output that is the result of multiple transformations. In regression problems, there is only one neuron in the output layer, which gives 1 output value. In classification problems, the number of neurons in the output layer is equal to the number of classes into which we want to divide the analysed objects.

The strength of artificial neural networks lies in the possibility of applying various activation functions that transform signals into the terms of the problem analysis (linear, s-shaped, Gaussian) and aggregation (linear, radial). Thanks to this, artificial neural networks are used in the analysis of problems where the algorithm for solving the problem is not fully known or must be frequently or quickly modified. They solve problems in the area of little-known phenomena and processes; the user does not have to declare any form of model in advance, and does not even have to be sure that any mathematical relationship can be modelled at all.

The network operation process consists of 3 stages:

- The learning stage, to adjust weights in order to obtain the most correct results. This stage takes place on the training set, extracted from the entire data set. It usually contains 50-80% of the observations.

- The testing stage, used to check the correct operation of the network on a new data set (approximately 10–30% of observations). If the correctness is not satisfactory, the network returns to the learning stage.
- The validation stage, in which final assessment regarding the correct operation of the network is made. Validation is performed on new data that was not used in the learning or testing stages. Most often, 10–30% of observations are allocated for validation.

3. Data and research procedure

Our dataset is formed using database of the Morningstar Direct for the local open-end investment fund. This local database is Poland. We find this market a representative setting from various reasons. In 1989 Poland started its period of transition from communism to capitalism and in 2004, 15 years later, together with a few other countries from Central Eastern Europe (CEE), Poland joined the European Union (EU). Since then, the country's gross domestic product (GDP) has always been positive—from a minimum of 1.4% to a maximum of 6.8% (see Table 1). In 2009, Poland was categorised by the World Bank as a mature economy and in 2018, by FTSE Russell, as a mature capital market. The Polish mutual fund market is the biggest amongst CEE countries and one of the most dynamically growing in the EU. On average, in 2005–2020, the open-end investment fund NAV has been growing there more than 20% p.a.; while, at the same time, on average in the EU it was 8.7%, and in the US, 6.5% (see Table 1). However, together with China, Poland is still among the countries with much lower ratios of stock market capitalisation to GDP than the most developed countries in the world; therefore, they tend to have fewer total net assets in regulated long-term funds that are relative to GDP (Investment Company Institute, 2020, see Figure 1.10). In other words, the open-end investment fund market in Poland is still much closer to that of an emerging market than it is to a mature one.

The Polish open-end investment funds are called “capital market funds” and they are offered mainly to retail investors. The style structure of the Polish open-end fund market is much more diversified than what is seen in the US or EU. Unlike in the US or EU, where taking NAV into account, the 2004–2018

Table 1. Macro and investment fund market data for USA, European Union and Poland

		2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	average p.a.
GDP																	
GDP (in %)	USA	3.8	3.5	2.9	1.9	-0.1	-2.5	2.6	1.6	2.2	1.8	2.5	2.9	1.6	2.2	2.9	2.0
	EU	2.4	2.3	2.2	3.1	0.5	-4.3	2.1	1.8	-0.4	0.3	1.8	2.3	2.0	2.5	2.0	1.2
	PL	5.3	3.5	5.8	6.8	4.2	2.8	3.6	5.0	1.6	1.4	3.3	3.8	3.1	4.8	5.1	4.0
Inflation																	
CPI (in %)	USA	3.3	3.4	2.5	4.1	0.1	2.7	1.5	3.0	1.7	1.5	0.8	0.7	2.1	2.1	1.9	2.1
	EU	2.4	2.3	2.2	3.2	2.2	1.5	2.7	3.0	2.3	1.0	-0.1	0.2	1.1	1.6	1.6	1.8
	PL	1.7	0.7	1.4	4.0	3.3	3.5	3.1	4.6	2.4	1.0	-1.0	-0.5	0.8	2.1	1.1	1.9
Stock market capitalisation as a percentage of GDP																	
	USA	133	130	141	138	79	105	115	101	116	144	151	138	147	166	148	130
	UK*	117	121	140	125	64	116	108	110	122	142	116	106	108	117	97	116
	PL	38	31	43	49	17	34	40	26	35	39	31	29	29	38	25	37
Open-end investment fund NAV growth																	
	USA	10.2%	8.8%	16.9%	15.4%	-20.0%	15.8%	6.3%	-1.6%	12.2%	15.3%	5.5%	-1.4%	4.4%	14.7%	-5.6%	6.5%
	EU	9.5%	22.8%	14.5%	7.5%	-22.9%	15.2%	14.0%	-1.4%	12.9%	9.5%	15.8%	19.5%	4.5%	11.6%	-3.2%	8.7%
	PL	36.8%	71.0%	62.3%	43.0%	-52.8%	28.2%	29.1%	-10.1%	38.2%	27.1%	7.5%	20.9%	-0.7%	13.8%	-10.6%	20.2%
Open-end investment fund structure																	
USA	equity	49.4%	53.7%	55.0%	56.1%	53.5%	38.0%	43.9%	47.3%	44.8%	45.5%	51.6%	52.4%	52.0%	52.4%	54.9%	50.2%
	money market	27.6%	23.5%	22.8%	22.5%	25.7%	39.8%	29.8%	23.7%	23.1%	20.6%	18.1%	17.2%	17.6%	16.7%	15.2%	22.6%
	bond	17.0%	16.0%	15.3%	14.4%	14.0%	16.3%	19.8%	21.9%	24.4%	26.0%	21.8%	21.8%	21.8%	22.3%	21.7%	19.8%
	multi-asset	6.0%	6.8%	7.0%	7.0%	6.8%	5.8%	6.5%	7.1%	7.6%	7.9%	8.5%	8.7%	8.6%	8.6%	8.2%	7.4%
EU	equity	35.0%	38.0%	41.0%	39.9%	29.1%	33.9%	36.0%	33.0%	33.0%	37.0%	38.0%	38.0%	37.0%	38.0%	39.0%	36.4%
	money market	21.0%	18.0%	16.0%	16.5%	25.8%	21.1%	20.0%	19.0%	16.0%	13.0%	13.0%	14.0%	13.0%	12.0%	2.0%	16.0%
	bond	27.0%	25.0%	23.0%	21.7%	22.9%	23.0%	23.0%	27.0%	29.0%	28.0%	28.0%	26.0%	27.0%	27.0%	23.0%	25.4%
	multi-asset	14.0%	13.0%	15.0%	15.5%	16.0%	16.4%	15.0%	16.0%	16.0%	16.0%	16.0%	17.0%	17.0%	18.0%	26.0%	16.5%
	other	3.0%	6.0%	5.0%	6.4%	6.3%	5.6%	6.0%	5.0%	6.0%	6.0%	5.0%	5.0%	6.0%	4.0%	10.0%	5.7%
PL	equity	12.7%	10.7%	19.5%	31.3%	23.4%	29.0%	27.7%	19.2%	16.9%	16.0%	13.8%	11.7%	10.9%	11.9%	13.4%	17.9%
	money market	14.4%	13.2%	8.2%	6.5%	10.3%	9.3%	13.2%	16.1%	9.5%	11.1%	13.5%	11.9%	12.4%	14.8%	21.9%	12.4%
	bond	21.2%	17.4%	7.2%	5.6%	14.6%	13.6%	14.4%	17.1%	28.3%	23.1%	20.1%	16.2%	16.8%	17.1%	17.4%	16.7%
	multi-asset	16.2%	18.6%	26.3%	31.7%	26.5%	32.7%	28.8%	18.9%	13.8%	11.3%	15.1%	12.8%	9.5%	11.7%	11.4%	19.0%
	other	35.5%	40.1%	38.8%	24.9%	25.2%	15.4%	15.9%	28.7%	31.5%	38.5%	37.5%	47.4%	50.4%	44.5%	35.9%	34.0%

* data for stock market capitalisation as a percentage of GDP of UK in 2004–2014 comes from the World Bank database and in 2015–2018 from the CEICdata.com.

Source: Eurostat, World Bank, Investment Company Institute, European Fund and Asset Management Association, Polish Chamber of Fund and Asset Management

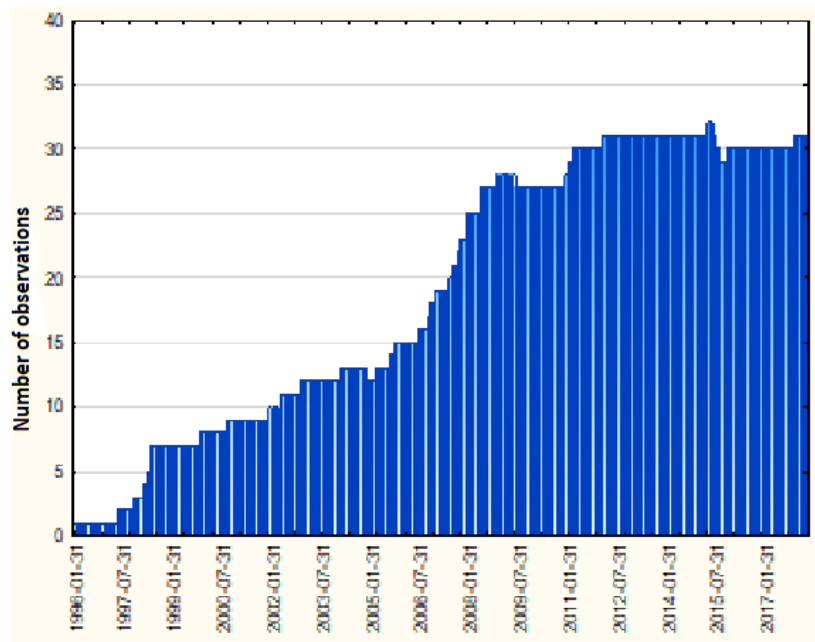


Figure 3. Number of observations in the studied period (12.1995-03.2018)
Source: own elaboration

equity funds counted for an average of 50.2% and 36.4%, respectively, in Poland, it was 17.9% (see Table 1). Polish multi-asset funds held, on average, 19%. Bond and money market funds accounted for 16.7% and 12.4%, respectively. This fund style structure suggests that the many Polish fund unitholders prefer to keep funds in their portfolios that are classified as moderate- or low-risk (at least lower than for their average American or European Union peers).

For the purposes of our study, we included data from the Morningstar Direct database on funds valued in PLN and funds belonging to large families of capital market funds registered in Poland, the net assets of which exceeded PLN 3 billion at the end of the research period. This condition was met by 12 out of 35 fund families, with a total market share of 87%. Ultimately, 37 equity funds were included in the sample (additionally, one fund was eliminated due to its short operation period of 2 months, thus yielding only 2 observations). The data on the funds start in December 1995 and go through March 2018, on a monthly basis. December 1995 was the first month of quoting the first equity fund operating on the Polish market and March 2018 was the last month of data availability. In total, the sample consisted of 4,645 observations (see Figure 3).

The following data were used as explanatory variables (**input variables**):

- logarithmic monthly return (R), calculated using quotations of the funds at the end of the given month,
- lagged logarithmic monthly returns by: 1 month $R(-1)$, 3 months $R(-3)$ and 6 months $R(-6)$, calculated using quotations of the funds at the end of the given month,
- logarithm of fund operation time counted in months (AGE),
- logarithm of fund size ($SIZE$, calculated on the basis of net asset value, NAV),
- the value of financial flow (cash flow, CF), calculated according to the following formula:

$$CF_{i,t} = \frac{NAV_{i,t} - NAV_{i,t-i}(1+R_{i,t})}{NAV_{i,t-i}},$$
- distribution method - dichotomous variable $DISTRIBUTION$ (0 - non-bank, 1-bank).

The descriptive statistics of these data are shown in Table 2. The diversity of returns and cash flow is very large but the values of AGE and $SIZE$ variable are homogeneous. The average value of monthly rate of return (R) is 0.334 with average variation of 5.561. The values range from -36.501 to 23.496. The descriptive statistics of lagged monthly returns are similar to those above. The average of variable AGE is 4.109 with standard deviation 0.975. As this variable is a logarithm of the numbers of months, the mean

Table 2. Descriptive statistics of quantitative independent variables

	N	Mean	Median	Minimum	Maximum	Std. Deviation
R	5055	0.33426	0.61367	-36.5010	23.49568	5.560999
R(-1)	5018	0.36612	0.64163	-36.5010	23.49568	5.566170
(R-3)	4944	0.37174	0.64123	-36.5010	23.49568	5.572185
(R-6)	4835	0.39409	0.66000	-36.5010	23.49568	5.613492
AGE	5055	4.10926	4.34381	0.0000	5.59099	0.975071
SIZE	5055	18.93488	19.09698	12.7528	22.53559	1.596097
CF	5054	0.05861	-0.00184	-0.3542	70.10703	1.094054

Source: own research

refers to approximately 5 years. The range of this variable before logarithmic transformation is from 1 month to over 22 years. The average of *SIZE* variable equals 18.935 and its standard deviation 1.596. The mean refers to value 167.230 Mio PLN, and the initial values of this variable range from 0.346 to 6124.666 Mio PLN. The mean value of cash flow is 0.059, its standard deviation equals 1.094, and the values are between -0.354 to 70.107. In the data sets 47.7% of observations are connected to equity funds distributed mainly via banks and 52.3% by the other channels.

Additionally, the following input variables were introduced into the analysis:

- *DATE* (month and year from which the observations come from),
- *FUND*—qualitative variable, the name of the fund that the data concern.

These variables were introduced because of the additional information they contain. We recognise that this information may prove important in the classification process. The variable *DATE* is a characteristic of the existing temporary economic situation. In turn, the *FUND* variable representing the name of a given fund may be treated as a synthetic variable, which is an indicator of the specificity regarding a given fund: the fund management model, the adopted strategy and its internal situation. The assumed investment risk may result not only from internal conditions; it may also be related to external factors—the current economic and/or market situation at a given moment.

The classification of funds was carried out on the basis of the following two **classifiers**, which are the adopted investment risk measures:

1. *BETA* ratio, which determines the degree of dependence regarding the fund's return rate and the return rate on the market portfolio, represented in our study by the WIG index.

According to the values of the beta coefficient, the observations were assigned to 3 classes:

$BETA < 0$,

$BETA (0, 1)$,

$BETA > 1$.

This is justified in the interpretation of the value of this indicator. When a fund's beta is greater than 1, the fund is said to be aggressive and has systematic risk greater than the market portfolio. A fund with a beta value of less than 1 and greater than 0 is defensive and reacts poorly to market changes. A fund with a beta below 0 is a fund whose returns are opposite to those of the market portfolio. This is a very rare but occurring case, thus it has been included. It should be noted that the *BETA* classes do not take into account the situation in which $BETA = 0$. Such a value of this ratio means that the fund's return does not respond to market changes, that is, the fund is risk-free. This was not the case during the period considered and therefore was not taken into account in our investigation.

According to these classification rules, there are 89.8% of all observations in class with *BETA* values between 0 and 1, 5.6% in class of negative *BETA* value and 4.6% in the third class – *BETA* more than 1. This uneven division is the result of only one style of funds being analysed. We expect a better distribution when more fund styles are taken into account in further research. The descriptive statistics of observations

Table 3. Descriptive statistics of quantitative independent variables in classes of BETA classifier

Variable	Class of BETA	N	Mean	Median	Minimum	Maximum	Std. Dev.
R	0	304	1.20025	1.30403	-14.628	20.05664	5.348625
	1	4522	0.40023	0.67158	-36.501	23.49568	5.441381
	2	229	-2.11814	-2.0185	-31.0289	20.50384	7.296892
R(-1)	0	300	1.10756	1.76688	-36.501	13.66871	5.552314
	1	4491	0.40151	0.63795	-34.4015	23.49568	5.483564
	2	227	-1.31399	-0.84962	-30.8209	14.91146	6.782666
(R-3)	0	292	1.82123	1.6856	-24.7961	21.52702	5.655938
	1	4426	0.32228	0.65432	-36.501	23.49568	5.533347
	2	226	-0.53241	-0.48742	-30.8209	12.79057	5.91571
(R-6)	0	278	0.98586	0.937	-24.6169	18.48296	5.07533
	1	4336	0.38388	0.66353	-36.501	23.49568	5.621471
	2	221	-0.14981	0.42433	-30.8209	12.49409	6.04421
AGE	0	304	3.50819	3.85015	0	5.12396	0.942185
	1	4522	4.13225	4.37574	0.6931	5.59099	0.961346
	2	229	4.45323	4.86753	0.6931	5.50533	0.979203
SIZE	0	304	17.75238	17.77637	12.7528	19.95648	1.344688
	1	4522	18.96870	19.12191	12.7671	22.53559	1.57812
	2	229	19.83668	19.83888	13.9245	22.51503	1.407123
CF	0	303	0.06953	0.01405	-0.3239	1.40754	0.189824
	1	4522	0.06084	-0.00249	-0.3542	70.10703	1.154969
	2	229	0.00008	-0.00713	-0.1938	2.24313	0.157846

Source: own research

in each class based on BETA classifier are shown in Table 3.

2. *SIGMA* – standard deviation of daily returns in individual months.

According to the *SIGMA* parameter, the observations were assigned to 2 classes:

$SIGMA < 0.75$,

$SIGMA > 0.75$.

The cut-off value of 0.75 corresponds to a variation of approximately 80%. This is a very high coefficient of variation; however, since we take into account equity funds, such high volatility is warranted. The *SIGMA* cut-off value of 0.75 divided the observations into two parts that were similar in number. If different styles

of mutual funds are considered, other limits of this classifier should also be taken into consideration.

According to those rules, there are 49.9% of all observations in class with *SIGMA* values less than 0.75 and 50.1% ones in the class with *SIGMA* more than 0.75. The descriptive statistics of observations in each class based on BETA classifier are shown in Table 4.

The **test procedure** is as follows: the *artificial neural network module* of the *Statistica 13* program was used for the calculations. The automatic network search procedure was applied. In this procedure, the user does not define the network parameters in advance; they are selected randomly. It is allowed to build both MLP and RBF networks. The MLP network, or multilayer perceptron, consists of neurons arranged in layers. It is a unidirectional network (that is, information always goes from input to output, and

Table 4. Descriptive statistics of quantitative independent variables in classes of SIGMA classifier

Variable	Class of SIGMA	N	Mean	Median	Minimum	Maximum	Std. Dev.
R	0	2480	1,33224	1,29754	-11,4557	13,83502	3,403976
	1	2575	-0,62690	-0,65290	-36,5010	23,49568	6,904846
R(-1)	0	2460	0,90702	0,81168	-16,7491	21,52702	3,267253
	1	2558	-0,15406	0,22143	-36,5010	23,49568	7,069031
(R-3)	0	2432	0,82333	0,89965	-18,8818	21,52702	3,877995
	1	2512	-0,06547	0,21742	-36,5010	23,49568	6,795085
(R-6)	0	2393	0,86726	0,91793	-16,7867	21,52702	4,400762
	1	2442	-0,06958	0,35119	-36,5010	23,49568	6,556710
Size	0	2480	4,25810	4,46591	0,0000	5,58350	0,903739
	1	2575	3,96591	4,18965	0,6931	5,59099	1,018972
WAN	0	2480	18,98222	19,03036	12,7528	22,38194	1,376150
	1	2575	18,88927	19,25083	13,0466	22,53559	1,781540
CF	0	2479	0,08046	-0,00222	-0,3542	70,10703	1,534043
	1	2575	0,03757	-0,00124	-0,3415	9,09067	0,288631

Source: own research

does not go back to earlier layers). The RBF network, that is, the network with radial base functions, has only 1 hidden layer and, according to the literature, it is better suited for solving classification problems (Broomhead & Lowe, 1988; Moody & Darken, 1989; Simon, 1994). The number of hidden neurons has been limited to the range of 5–20 for the MLP network and 10–30 for the RBF network. Various activation functions (linear, logistic, exponential, hyperbolic tangent) have been adopted.

From all possible combinations of parameters, 50 networks were selected, of which the best 5 (providing the best classification correctness) were selected for the final classification.

The classification was made on the basis of the network set – that is, the observations of a given fund were allocated to the class indicated by the largest number of networks.

4. Results

4.1. Classification based on *BETA* classifier

As a result of building ANN procedure, there are 45 neurons in the input layer and 3 neurons in the output

layer for each out of 5 of the best network. The number of neurons in the hidden layer vary from 6 to 13. All finally chosen networks are multilayer perceptrons with linear aggregation functions. The activation functions are logistics sigmoid in two networks, exponential in two networks and one hyperbolic tangent in hidden layer neurons. In the output layer there are hyperbolic tangents in three cases, logistics sigmoid and softmax (see Table 5).

The adoption of the values 0 and 1 as the classifier's cut-off values was substantively justified; however, it resulted in the defensive funds class ($BETA = < 0; 1 >$) containing as much as 89.8% of all observations. This means that if artificial neural networks were not used to assess the investment risk taken by the funds, but in advance it was assumed that all observations belong to the $BETA$ class = $< 0; 1 >$, the correctness of our classification would be high and would amount to almost 90%. Within this context, the correctness of the network operation presented in the classification matrix (Table 6) should be assessed.

Based on the values of the input variables, neural networks correctly classified 4,189 observations (90.2%). There was a significant improvement in the quality of the classification within the class of defensive funds, where the networks correctly classified 99.4% of the observations. The percentage of correct classifications in the $BETA < 0$ class and in

Table 5. Description of the network set used in classification based on BETA classifier

ANN No.	ANN type	Number of hidden neurons	Activation function (hidden layer)	Activation function (output layer)
1	MLP	13	Hyperbolic tangent	Hyperbolic tangent
2	MLP	6	Logistic sigmoid	Softmax
3	MPL	11	Exponential	Hyperbolic tangent
4	MLP	13	Exponential	Logistic sigmoid
5	MLP	15	Logistic sigmoid	Hyperbolic tangent

Source: own research

Table 6. Classification matrix based on BETA classifier

Observed class	Percentage of correct classification	Predicted class			TOTAL
		<0	<0 ; 1>	> 1	
< 0	14%	36	225	0	261
<0 ; 1>	99%	23	4143	3	4169
> 1	5%	0	205	10	215
TOTAL	90%	59	4573	13	4645

Source: own research

the aggressive fund class ($BETA > 1$) is very low and amounts to 14% and 5%, respectively. Moreover, the range of $BETA$ parameter values in the dominant class is quite wide, which is why it concerns the majority of observations.

The results show that many of the surveyed equity funds were defensive in their nature. After equity funds, we should expect an aggressive rather than defensive investment policy. This perhaps resulted from a very long research horizon, which smoothed out the volatility regarding the returns of these funds over time. This could be overlapped with the defensive actions of managers (let us bear in mind that they are from the companies managing the oldest funds of the largest size), who prefer allocating the funds' assets in the value stocks that are less volatile than the growth stocks. Such approach to management meets the preferences of Polish individual investors—the structure of the fund market shows that in general they are rather risk averse, not risk takers.

4.2. Classification based on SIGMA classifier

The procedure of building of ANN resulted with the set consisting of 5 networks of MLP type (the aggregation function is linear). There are 45 neurons in the input layer and 2 neurons in the output layer. The number of neurons in the hidden layer varies from 6 to 13. In the hidden layer there is logistics sigmoid in three networks as well as the exponential function and the hyperbolic tangent in the two remaining ones. In output layers the softmax appears three times and the hyperbolic tangent twice. The final set of the neural network used in this classification can be seen in Table 7.

The cut-off value of the *SIGMA* classifier divided the observations into 2 classes. The *SIGMA* parameter value < 0.75 occurred in 49.9% of observations, while the remaining 50.1% of observations had a *SIGMA* value > 0.75 . This means that if all observations (without the use of neural networks) were classified into a more numerous class, a correctness of about 50% would be obtained.

Table 7. Description of the network set used in classification based on SIGMA classifier

ANN No.	ANN type	Number of hidden neurons	Activation function (hidden layer)	Activation function (output layer)
1	MLP	6	Logistic sigmoid	Softmax
2	MLP	9	Exponential	Hyperbolic tangent
3	MPL	7	Hyperbolic tangent	Softmax
4	MLP	13	Logistic sigmoid	Hyperbolic tangent
5	MLP	9	Logistic sigmoid	Softmax

Source: own research

Table 8. Classification matrix based on SIGMA classifier

Observed class	Percentage of correct classification	Predicted class		TOTAL
		< 0.75	> 0.75	
< 0.75	82.9%	1921	395	2316
> 0.75	76.7%	543	1786	2329
TOTAL	79.8%	2464	2181	4645

Source: own research.

The observations classified in the first group (*SIGMA* class < 0.75) mostly belong to the funds distributed by the bank channel, characterised by a longer average age on the market, a slightly lower value of net assets and a higher average monthly rate of return. On the other hand, the observations classified to the *SIGMA* > 0.75 group are mostly “non-bank” funds. In this group the fund age, size and monthly returns are more varied compared to the observations from the first group.

The classification made with the use of artificial neural networks provides 79.8% of correct classifications (Table 8). The quality of the network operation can be considered to be satisfactory. This means that the assumed input variables that were the fund characteristics can be used to determine the fund’s operating style and, to some extent, replace the standard deviation as a commonly accepted measure of risk.

The correctness of the classification in the group of observations for which the *SIGMA* value < 0.75 is 82.9%, in the group of observations with the *SIGMA* value > 0.75, the correctness of the classification is slightly lower and amounts to 76.7%.

The results confirm the legitimacy of using artificial neural networks as a tool for classifying

equity investment funds. Overall, standard deviation turns out to be a better classifier than the beta ratio. Standard deviation is the most recognised measure of risk that is published in the Key Information for Investors Document (KIID) obligatory for open-end investment funds operating in the European Union. Therefore our results confirm that it can be used in practice by investors as an indicator of the open-end investment funds classification.

4.3. Other characteristics essential to fund classification

Artificial neural networks are a method of analysis known as a ‘black box’, which means that we enter input and get the result of the network, but we cannot fully control or analyse the rules producing the result or the rules that the network has detected. Therefore, the interpretation of the relationship between the characteristics of the funds used in the study as input variables, and the applied risk measures, is not easy and obvious. Part of this interpretation is possible thanks to the global sensitivity analysis, which allows comparison of the quality of the network classification with the given variable for each input variable used in the network and after its removal. A value higher than 1 obtained in such an analysis proves that the network works better with this variable than without it—thus, it is important for the quality of the classification. It is related to the investment risk taken by the funds. A value close to 1 proves that the network works just as well with this variable as without it—this variable adds nothing to the classification. A value lower than 1 indicates that the network without this variable provides better correctness of the classification than with this variable—it is a variable disturbing the classification.

Table 9. Global analysis of network sensitivity (*SIGMA* classifier)

DISTRIBUTION	FUND	DATE	AGE	R(-1)	SIZE	R	R(-6)	R(-3)	CF
2.90	2.75	1.51	1.35	1.20	1.19	1.12	1.03	1.02	1.01

Source: own research.

The mean results of the overall sensitivity analysis obtained for the network group we investigated are presented in Table 9. The sensitivity analysis shows that the characteristics important for the fund classification based on the sigma risk taken by funds are primarily the distribution channel, the name of the fund and the current economic situation. Especially the last characteristic has not yet been considered and found important by the literature in the classification of open-end investment funds. Qureshi et al. (2019) prove that investors choose to invest in equity-based funds when economic conditions are good, but when they are poor, investors move to the income-type funds. Therefore the fund flows forecast the economic conditions (Ferson & Kim, 2012; Jank, 2012; Kopsch et al., 2015). We find the other side of the fund-macroeconomy relation—the economic conditions may be considered as the open-end fund classification guide—depending on the business cycle, the equity funds may be more or less tempted to drift the style in order to rescue the fund flows balance.

Slightly less crucial, but still important for the classification, are the age and the size of the fund, as well as the current and last-month return. The less important variables turned out to be the 3 and 6 month lagged returns and the cash flows of the funds. Investors should pay attention to these characteristics when they make investment decisions about the funds. Presumably older and bigger funds with more stable short-term returns follow the classification reflected in their names more than funds that are younger, smaller and more volatile. It does not mean investors should avoid the latter funds. They should simply be aware such funds may be more exposed to the style drift.

Our results confirm the findings of Cao et al. (2017) in the case of the fund size, and of Herrmann et al. (2016) in the case of the one-month lagged return. The other results concerning other characteristics that we found essential to the classification of open-end investment funds based on the standard deviation seem to be pioneering. They should be further investigated in expanded research with the same (or similar) methodology in order to support their meaning for the investment fund style analysis.

5. Conclusions

The need for in-depth diagnosis of the classification of open-end mutual funds, of which so many are on offer nowadays, has been recently increasing mainly for two reasons. First, financial advisors have increased responsibility to offer funds objectively tailored to the capabilities of an individual investor. This investor has little knowledge of the financial market and has to invest her savings there because the current pension systems are not able to provide her with funds for old age. Second, the need to reduce the costs of management and distribution of actively managed funds, which, for example, is imposed by a competitive market in the US, and the MiFID II Directive in the EU, means that accurate classification is needed. New technologies come with help, including machine learning tools, which, in this article, were used to verify the correctness of the classification of equity funds managed by the largest investment fund companies in an emerging fund market in Europe. Classification based on artificial neural networks allowed us to partially confirm the standard class identification of the researched funds. Artificial networks provided more applicable results in the case of standard deviation as a classifier, where the classification was evenly distributed over 2 groups and was correct in about 87%. This result is quite similar to that achieved by Białous and Truszkowski (2009). In the case of the *BETA* coefficient, the classification accuracy was 99%. However, almost all the observations were in the first class, that is in the class with the *BETA* coefficient ranging from 0 to 1. This means that the analysed funds were defensive at the time. This conclusion is consistent with the results of an early study on the values of beta ratio of equity funds in Poland by Sarnowski (2003). Our results are quite surprising, though, due to the fact that we are dealing with a much longer time series analysed, and within the setting of equity funds that have been actively managed with many years' exposure to international markets (certainly since Poland's accession to the EU in 2004). The European Commission report of 2018 indicates that managers

of open investment funds in Poland charge fees that are twice as high as those in other EU countries (see European Commission, (2018). If so, managed equity funds should be aggressive, with the values of their beta above one ($\beta > 1$), not lower than one. The issue of classifying investment funds in the context of their management activity level is beyond the scope of this article; however, this conclusion is a motivation for further research on the factors behind this situation.

The obtained results are sufficient to confirm the validity of using machine learning as a tool for classification (and further, for grouping) of open-end investment funds. To bolster the credibility of these results, further research should expand the sample to include other styles of funds in Polish and other markets, including the emerging fund market of China or the mature fund markets of Europe or the USA. Other risk measures, such as the Synthetic Risk and Reward Indicator (SSRI) published in the Key Information for Investors may be investigated as a risk measure as well. It is also worth considering other features of funds, such as those characterising the investors, the teams managing funds, or those describing the entire family of funds.

Acknowledgments

The project was financed within the Regional Initiative for Excellence programme of the Minister of Science and Higher Education of Poland, years 2019-2022, grant no. 004/RID/2018/19, financing 3,000,000 PLN.

References

- Bams, D., Otten, R., & Ramezanifar, E. (2015). Mutual Fund Objective Misclassification: Causes and Consequences. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2648375>
- Białous, K., & Truszkowski, J. (2009). Analiza stylu inwestowania polskich funduszy inwestycyjnych. *Ruch Prawniczy, Ekonomiczny i Socjologiczny*, LXXI(4), 195–223.
- Broomhead, B. S., & Lowe, D. (1988). Multivariable Functional Interpolation and Adaptive Networks. In *Complex Systems 2* (pp. 321–355). https://www.complex-systems.com/abstracts/v02_i03_a05/
- Brown, K. C., Harlow, W. V., & Zhang, H. (2015). Investment Style Volatility and Mutual Fund Performance. *Working Paper*. <http://faculty.mcombs.utexas.edu/keith.brown/Research/stylevolatility-wp.pdf>
- Brown, S. J., & Goetzmann, W. N. (1997). Mutual fund styles. *Journal of Financial Economics*, 43(3), 373–399. [https://doi.org/10.1016/S0304-405X\(96\)00898-7](https://doi.org/10.1016/S0304-405X(96)00898-7)
- Cao, C., Iliev, P., & Velthuis, R. (2017). Style drift: Evidence from small-cap mutual funds. *Journal of Banking and Finance*, 78, 42–57. <https://doi.org/10.1016/j.jbankfin.2017.01.009>
- Castellanos, A. R., & Alonso, B. V. (2004). Spanish Mutual Fund Misclassification. *The Journal of Investing*, 14(1), 41–51. https://www.academia.edu/24331925/Spanish_Mutual_Fund_Misclassification
- Chiang, W. C., Urban, T. L., & Baldrige, G. W. (1996). A neural network approach to mutual fund net asset value forecasting. *Omega*, 24(2), 205–215. [https://doi.org/10.1016/0305-0483\(95\)00059-3](https://doi.org/10.1016/0305-0483(95)00059-3)
- Chua, A. K. P., & Tam, O. K. (2020). The shrouded business of style drift in active mutual funds. *Journal of Corporate Finance*, 64, 101667. <https://doi.org/10.1016/j.jcorpfin.2020.101667>
- Cremers, K. J. M., & Petajisto, A. (2009). How Active Is Your Fund Manager A New Measure That Predicts Performance. *Review of Financial Studies*, 22(9), 3329–3365. <https://doi.org/10.1093/rfs/hhp057>
- Daniel, K., Grinblatt, M., Titman, S., & Wermers, R. (1997). Measuring Mutual Fund Performance with Characteristic-Based Benchmarks. *The Journal of Finance*, 52(3), 1035–1058. <https://doi.org/10.1111/j.1540-6261.1997.tb02724.x>
- DiBartolomeo, D., & Witkowski, E. (1997). Mutual fund misclassification: Evidence based on style analysis. *Financial Analysts Journal*, 53(5), 32–43. <https://doi.org/10.2469/faj.v53.n5.2115>
- European Commission. (2018). *Study on the distribution systems of retail investment products*. https://ec.europa.eu/info/publications/180425-retail-investment-products-distribution-systems_en
- Person, W. E., & Kim, M. S. (2012). The factor structure of mutual fund flows. *International Journal of Portfolio Analysis and Management*, 1(2), 112. <https://doi.org/10.1504/ijpam.2012.049214>
- Gandhmal, D. P., & Kumar, K. (2019). Systematic analysis and review of stock market prediction

techniques. *Computer Science Review*, 34, 100190. <https://doi.org/10.1016/j.cosrev.2019.08.001>

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical Asset Pricing via Machine Learning. *Review of Financial Studies*, 33(5), 2223–2273. <https://doi.org/10.1093/rfs/hhaa009>

Herrmann, U., Rohleder, M., & Scholz, H. (2016). Does style-shifting activity predict performance? Evidence from equity mutual funds. *Quarterly Review of Economics and Finance*, 59, 112–130. <https://doi.org/10.1016/j.qref.2015.03.003>

Indro, D. C., Jiang, C. X., Patuwo, B. E., & Zhang, G. P. (1999). Predicting mutual fund performance using artificial neural networks. *Omega*, 27(3), 373–380. [https://doi.org/10.1016/S0305-0483\(98\)00048-6](https://doi.org/10.1016/S0305-0483(98)00048-6)

Investment Company Institute. (2020). *2020 ICI Fact Book*. <https://www.ici.org>

Jank, S. (2012). Mutual fund flows, expected returns, and the real economy. *Journal of Banking and Finance*, 36(11), 3060–3070. <https://doi.org/10.1016/j.jbankfin.2012.07.004>

Jiang, J., Liao, L., Wang, Z., & Xiang, H. (2020). Financial literacy and retail investors' financial welfare: Evidence from mutual fund investment outcomes in China. *Pacific Basin Finance Journal*, 59 (November), 101242. <https://doi.org/10.1016/j.pacfin.2019.101242>

Jung, D., Glaser, F., & Köpplin, W. (2019). Robo-advisory: Opportunities and risks for the future of financial advisory. In *Contributions to Management Science* (pp. 405–427). Springer. https://doi.org/10.1007/978-3-319-95999-3_20

Kim, M., Shukla, R., & Tomas, M. (2000). Mutual fund objective misclassification. *Journal of Economics and Business*, 52(4), 309–323. [https://doi.org/10.1016/S0148-6195\(00\)00022-9](https://doi.org/10.1016/S0148-6195(00)00022-9)

Kim, T. H., White, H., & Stone, D. (2005). Asymptotic and Bayesian confidence intervals for sharpe-style weights. *Journal of Financial Econometrics*, 3(3), 315–343. <https://doi.org/10.1093/jfinec/nbi015>

Kopsch, F., Song, H. S., & Wilhelmsson, M. (2015). Determinants of mutual fund flows. *Managerial Finance*, 41(1), 10–25. <https://doi.org/10.1108/MF-06-2013-0161>

Mason, A., McGroarty, F., & Thomas, S. (2012). Style analysis for diversified US equity funds. *Journal of Asset Management*, 13(3), 170–185. <https://doi.org/10.1057/jam.2012.6>

Mohanti, D., & Priyan, P. K. (2018). Style-exposure analysis of large-cap equity mutual funds in India. *IIMB Management Review*, 30(3), 219–228. <https://doi.org/10.1016/j.iimb.2018.01.010>

Moody, J., & Darken, C. J. (1989). Fast Learning in Networks of Locally-Tuned Processing Units. *Neural Computation*, 1(2), 281–294. <https://doi.org/10.1162/neco.1989.1.2.281>

Moreno, D., Marco, P., & Olmeda, I. (2006). Self-organizing maps could improve the classification of Spanish mutual funds. *European Journal of Operational Research*, 174(2), 1039–1054. <https://ideas.repec.org/a/eee/ejores/v174y2006i2p1039-1054.html>

Müller, S., & Weber, M. (2010). Financial Literacy and Mutual Fund Investments: Who Buys Actively Managed Funds? *Schmalenbach Business Review*, 62(2), 126–153. <https://doi.org/10.1007/bf03396802>

Pan, W. T., Han, S. Z., Yang, H. L., & Chen, X. Y. (2019). Prediction of mutual fund net value based on data mining model. *Cluster Computing*, 22, 9455–9460. <https://doi.org/10.1007/s10586-018-2272-2>

Pandurang, G. D., & Kumar, K. (2019, February 1). Ensemble Computations on Stock Market: A Standardized Review for Future Directions. *Proceedings of 2019 3rd IEEE International Conference on Electrical, Computer and Communication Technologies, ICECCT 2019*. <https://doi.org/10.1109/ICECCT.2019.8869158>

Qureshi, F., Khan, H. H., Rehman, I. U., Ghafoor, A., & Qureshi, S. (2019). Mutual fund flows and investors' expectations in BRICS economies: Implications for international diversification. *Economic Systems*, 43(1), 130–150. <https://doi.org/10.1016/j.ecosys.2018.09.003>

Rout, M., Koudjonou, K. M., & Satapathy, S. C. (2020). Analysis of net asset value prediction using low complexity neural network with various expansion techniques. *Evolutionary Intelligence*, 0123456789. <https://doi.org/10.1007/s12065-020-00365-0>

Sarnowski, K. (2003). Kłasyfikacja funduszy inwestycyjnych za pomocą jednoczynnikowej analizy wariancji. *Prace Naukowe Akademii Ekonomicznej We Wrocławiu, nr 991 Inwestycje finansowe i ubezpieczenia-tendencje światowe a polski rynek*, 568–574.

Sensoy, B. A. (2009). Performance evaluation and self-designated benchmark indexes in the mutual fund industry. *Journal of Financial Economics*, 92(1), 25–39. <https://doi.org/10.1016/j.jfineco.2008.02.011>

Sha, Y. (2020). The devil in the style: Mutual fund style drift, performance and common risk factors.

Economic Modelling, 86 (October), 264–273. <https://doi.org/10.1016/j.econmod.2019.10.004>

Sharpe, W. (1992). Asset allocation: Management style and performance measurement. *The Journal of Portfolio Management*, 18(2), 7–19. <https://doi.org/10.3905/jpm.1992.409394>

Simon, H. (1994). *Neural Networks: A Comprehensive Foundation*. Macmillan Publishing.

Wang, K., & Huang, S. (2010). Using fast adaptive neural network classifier for mutual fund performance evaluation. *Expert Systems with Applications*, 37(8), 6007–6011. <https://doi.org/10.1016/j.eswa.2010.02.003>

Wermers, R. R. (2012). Matter of Style: The Causes and Consequences of Style Drift in Institutional Portfolios. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2024259>

Zamojska-Adamczak, A. (2005). *Style inwestowania jako kryterium klasyfikacji funduszy inwestycyjnych* (Investment styles as a classification criterion for investment funds).

Zhou, S., Cheng, Y., & Zhou, X. (2018). Research on the “Investment Style Drift” Phenomenon of China’s Open-end Equity Funds. *Management World*, 6, 175–176.