

FINDING OPPORTUNITY WINDOWS IN TIME SERIES DATA USING THE SLIDING WINDOW TECHNIQUE: THE CASE OF STOCK EXCHANGES

Necmi Gürsakal

Fenerbahçe University, Faculty of Engineering and Architecture, Istanbul, Turkey

e-mail: necmi.gursakal@fbu.edu.tr

ORCID: 0000-0002-7909-3734

Fırat Melih Yılmaz

Dokuz Eylül University, Institute of Social Sciences, Izmir, Turkey

e-mail: firatmelihyilmaz@outlook.com

ORCID: 0000-0001-5816-4321

Erginbay Uğurlu

Istanbul Aydın University, Department of International Trade, Istanbul, Turkey

e-mail: erginbayugurlu@aydin.edu.tr

ORCID: 0000-0002-1297-1993

© 2020 Necmi Gürsakal, Fırat Melih Yılmaz, Erginbay Uğurlu

This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License.

To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/4.0/>

Quote as: Gürsakal, N., Yılmaz, F. M., and Uğurlu, E. (2020). Finding opportunity windows in time series data using the sliding window technique: The case of stock exchanges. *Econometrics. Ekonometria. Advances in Applied Data Analysis*, 24(3).

DOI: 10.15611/ead.2020.3.01

JEL Classification: C22, C53, C58

Abstract: Data have shapes, and human intelligence and perception have to classify the forms of data to understand and interpret them. This article uses a sliding window technique and the main aim is to answer two questions. Is there an opportunity window in time series of stock exchange index? The second question is how to find a way to use the opportunity window if there is one. The authors defined the term opportunity window as a window that is generated in the sliding window technique and can be used for forecasting. In analysis, the study determined the different frequencies and explained how to evaluate opportunity windows embedded using time series data for the S&P 500, the DJIA, and the Russell 2000 indices. As a result, for the S&P 500 the last days of the patterns 0111, 1100, 0011; for the DJIA the last days of the patterns 0101, 1001, 0011; and finally for the Russell 2000, the last days of the patterns 0100, 1001, 1100 are opportunity windows for prediction.

Keywords: time series, data science, patterns, sliding window.

1. Introduction

Data science is a pragmatic and goal-driven method that extracts value from data. There are many different ways to analyse the data. By developing computer technologies, Machine Learning (ML) methods have been proposed to analyse data. Machine learning is used to discover knowledge for making intelligent decisions. Although some authors are optimistic about ML methods, Makridakis, Spiliotis, and Assimakopoulos (2018) state that ML methods can be applied to forecasting¹. In the paper, the authors aimed to show that traditional statistical methods are more accurate than ML ones, and searched for the reasons and found ways to reverse the situation. Moreover, Gareth, Witten, Hastie, and Tibshirani (2017) argued that stock movements and their results show that 56% of the daily movements were predicted correctly.

Data science is a very suitable area for solving a business problem using data in an interdisciplinary approach and turning this solution into money. Forecast and prediction, in general, are related to quantitative values but it should also be accepted that qualitative forecasts such as combinations of up and down for a time period and predictions such as up or down for a day in stock markets can be useful.

In general, it is more common to try to transform nominal data into interval data, but on the contrary, the authors did the exact opposite and transformed interval data into nominal data. If one only wants to predict whether the stock market index will move upwards or downwards for the next day, the probability of these outcomes will be 0.50 and 0.50, due to the lack of knowledge. Let us suppose that we have a ‘modest’ purpose and want to use a nominal scale instead of an interval scale. Instead of using the opening or closing values of a stock market index or stock values, one can transform these data into a nominal (categorical) scale by taking these values as up and down. The purpose of this transformation is to create some up and down patterns. Although the technique is to predict a pattern, the study aimed to find forecast opportunities by using patterns. If there was the opportunity to forecast using sliding windows, the window which was used for the forecast was called the opportunity window.

Using up and down pattern data, an opportunity window is a proper sliding window length that creates statistically significant differences in the observed proportions of patterns for the prediction of the future direction of the index. Only at these points of time, are the proportions of the up and down movements significantly different from each other and this can be an opportunity to predict the direction of an index.

For the financial data, various time series analysis techniques were widely used. They were used for determining future prices of stocks, modelling prices or returns

¹ They wrote that “There is no reason that the same type of breakthroughs cannot be achieved with ML methods applied to forecasting”.

of stocks, and modelling affecting factors of stocks etc. In this paper, the authors aimed to see if there is an opportunity window in time series data using the sliding window technique. The study used three different stock exchange indices as a time series data, namely the S&P 500, the DJIA, and the Russell 2000.

The contribution of this paper to the literature is that it used the sliding window technique to stock exchanges indices. The most important contribution is adding the term ‘opportunity window’ to the literature. The rest of the paper is organized as follows. Section 2 presents a literature review. Section 3 gives the description of the steps required for the sliding window technique. Section 4 presents the data and the results of the empirical application. Finally, the paper is concluded in Section 5.

2. Literature review

Two approaches can be selected when working with time series. In the first, the time series are viewed from a general window as a whole, where the units of analysis are daily, monthly, and annual values. In a second, instead of looking at a general window, the time series are viewed from many sliding windows where the analysis units consist of as much data as the window width covers.

There are existing systems such as Support Vector Machine (SVM) and Association Rule Mining (ARM) to implement time series data, but these systems have some issues with optimal rules discovery and segmentation on time series data, as well as lower accuracy and higher complexity (Senthil and Suseendran, 2018). To solve this issue Sliding Window Techniques (SWT) based systems were used. Two systems are mentioned in Senthil and Suseendran (2018), i.e. Improved ARM with Enhanced SVM (SWT-IARM with ESVM). This paper used the sliding vector time series analysis (SWTS). By using the SWTS, the resulting predictor can be used at the heart of the powertrain predictive controller (Mozaffari, Mozaffari, and Azad, 2015). When using SWTS, the data in the analysis must be presented in a state-space format. SWTS split the database into several finite-length segments and tries to relate z past data to the p ahead data. There are different kinds of Windows models, therefore to understand them, firstly data stream monitoring should be defined.

The categories of data stream monitoring are point monitoring and aggregate monitoring (Zhu and Shasha, 2003). The sliding windows model is in the aggregate monitoring category in which three well-known models are used, namely landmark windows, sliding windows, and damped windows. Zhu and Shasha (2003) state that the sliding window is most widely used in data stream monitoring. This method can be used in different kinds of fields. For example in computerized analysis of medical data, signals in the form of time series are frequently used for diagnostic and prognostic purposes data sets preprocessed and instances created and analyzed. Öztürk Katircioğlu, Güvenir, Ravens, and Baykal (2017) is a paper in the medical field, while Yahmed et al. (2015) is a paper regarding meteorology. Öztürk et al. (2017) proposed a different method which they calledled the Window-based Time Series

Feature Extraction (WTC) method. The authors proposed a new term, *shapelets*, which is a transformation technique. A shapelet is a time series subsequence identified as representative of a certain dataset and is essentially a subsequence that is extracted from the time series instances. Yahmed, Azuraliza, RazakHamdan, Almahdi, and Abdullah (2015) used sliding windows for weather forecasts, whereas Yu, Zhu, Li, and Wan (2014) to detect outliers in hydrological time series data.

As computer and Internet technologies have evolved, time series have been transformed into stream data and the size of these data has increased rapidly. Rajalakshmi and Dinakaran (2015) presented a comparison of previous studies on time series data using pattern-matching based on their algorithms, and the advantages and disadvantages of such studies. In the paper, fields of the stream time-series data are mentioned which include Internet traffic analysis, sensor network monitoring, moving object search, and financial data analysis. These kinds of analyses need constant monitoring of stream time series.

Similarly to this paper, it is used for stock index data. Hota, Handa, and Shrivastava (2017) applied Radial Basis Function Network (RBFN) for data prediction using the concept of a sliding window. The authors argue that although different intelligent techniques are used for time series analysis, the prediction of next day closing price on the basis of current day price is not appropriate.

The sliding window technique can be used for this prediction. It is based on the accumulation of historical time series data with segmentation. In the process, an approximation is carried out temporarily over actual value and size of the window and segment increases until the less error approximation is reached (Mozaffari, Mozaffari, and Azad, 2015; Vafaeipour, Rahbari, Rosen, Fazelpour, and Ansarirad, 2014; Hota, Handa, and Shrivastava, 2017).

3. Methodology

Ali, Jones, Xie, and Williams (2019) present logic which is very close to this article's application. The authors assert that to understand the time series data some features need to be determined such as changes over time, repeated patterns, outliers, and labeling data instances. As the authors used a univariate time series, Figure 1 presents a summary of the univariate methodology by Ali et al. (2019). The methodology is based on anomalous and clusters patterns which show outlier patterns and repeated patterns, respectively. The aim was to explore and interpret these patterns.

The process in Ali et al. (2019) starts with collecting raw data, and after several steps the last one is reached, which allows users to interact with the system and change the parameters, which in turn helps to improve and fit the users tasks. Raw data are collected and then transformed into 2D space passing through a multi-step process. Through this transformation time series, data arrive in the high-dimensional space. The transformation can be carried out based on the type of time-dependent variables. If the time-dependent variable is univariate PCA, t-SNE, or UMAP, they

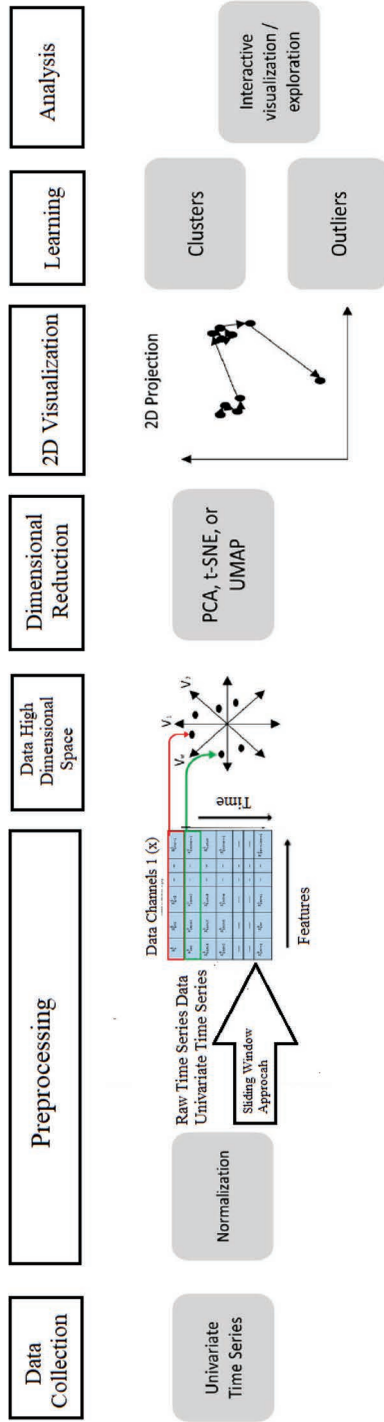


Fig. 1. An overview of Ali et al. (2019) proposal of visual analytics approach

Source: (Ali et al., 2019, p. 1016).

are applied directly on the sliding window matrix. However, if the time-dependent variable is multivariate, the first step DCAE is applied to extract important features, which are then projected into 2D space using PCA, t-SNA, or UMAP.

In machine learning, it is important to define the *pattern*. Sverdlov (2015) defines a pattern as some structure in concept space. In the machine-learning process, attributes define points in concept space, and space is generally a region or subspace. Moreover, it can be a simple point or N-dimensional object, and can also be defined in numeric terms or non-numeric attributes.

In this paper, we use stock market indices as they are examples of interval type data. In general, it is preferred to use the data obtained with high-level scales in statistical applications. However, when estimating at a stock exchange, it is important to predict that the value of the index or stock will tend to increase or decrease the next day. If one wants to make a near-perfect prediction, it would be useful to add to this trend the amount of increase or decrease. However, it is difficult to obtain such detailed estimates, and these estimates are mostly forecasted as a range. Therefore, the authors made an application by converting the scale of the financial time series into a low nominal, categorical scale in order to calculate the probabilities of these categorically defined patterns in the past and then made predictions for these probabilities using R.

Table 1. Number of pattern types

Pattern length (n)	Number of pattern types
2	4
3	8
4	16
5	32
6	64
7	128
8	256
9	512
10	1024

Source: authors' calculation.

A univariate time series $X = \{x_t \mid 1 \leq t \leq i\}$ is a sequence of the observations vector and mostly these observations are collected at equally spaced time intervals. If the length of window n is taken, one can determine sliding windows segments overlapping or non-overlapping. In Figure 2, one see an overlapping sliding window when n is taken as 5.

Omitting the *NO CHANGE* result in two stock exchange time series with only *UP and DOWN* categories for stock exchanges, Table 1 shows how the number of pattern types increased in line with the pattern length increases. In any case complexity increases for analysis as the number of pattern types increases and goes beyond common recognition. However, it should be noted that machines do not forget anything and do not mix them up with these patterns as humans do.

If we have $n = 2$, we may have only UU, UD, DU, and DD results. If $n = 3$ we will have UUU, UUD, UDU, UDD, DUU, DUD, DDU, DDD results for the next day. In this application, Down is shown by 0 and Up by 1.

A univariate time series $X = \{x_t \mid 1 \leq t \leq i\}$ is a sequence of the observations vector and mostly

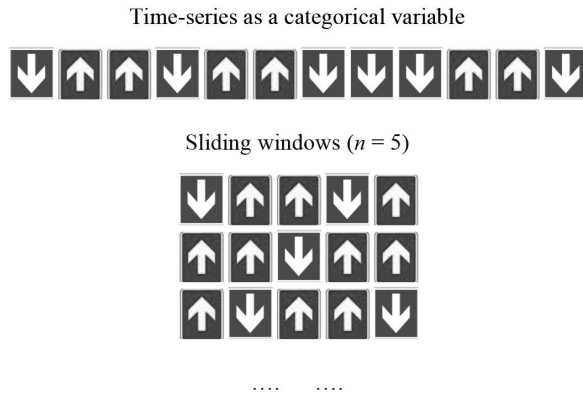


Fig. 2. Process of sliding window

Source: authors’ work.

4. Data and empirical application

In this application for finding opportunity windows in time series of stock exchange indices using the sliding window technique, the authors used three stock exchange indices: the S&P 500, the Russell 2000, and the Dow Jones Industrial average. All of them measure the price movements of stocks of the US economy and use different sets of stocks (holdings) and weighting methodologies.

The S&P 500 Index is the most widely-used benchmark for the large capitalization of U-based stocks. The index was compiled in 1957 and comprises 500 stocks. The firms are updated according to selected Standard and Poor’s criteria. Each year approximately 20 firms are added and 20 firms are dropped from the index (Siegel and Schwartz, 2006). Hora and Jalbert (2006) state that the Dow Jones Industrial Average (DJIA) is “the single most important indicator of the health and direction of US capital markets.” The DJIA lists 30 US blue-chip companies (spglobal.com, 2020); it comprise 12 companies in 1986, and of these 12 original companies, only General Electric remained in 2001 (Lerman, 2001). Transportation and utilities industries are not covered in the DJIA. The Russell 2000 index can be used to measure small-cap equity funds’ movements (Cai and Houge, 2008). The index was created in 1984 by the Frank Russell Company.

Figure 3 shows the daily data of the S&P 500 index for the period of 01.01.2015-28.08.2019.

Figures 4 and 5 present the time series graph of the DJIA and the Russell 2000 index, respectively. These two graphs were drawn for the same time period as Figure 3.



Fig. 3. Closing values of the S&P500

Source: (Yahoo Finance, n.d.).



Fig. 4. Closing values of the DJIA

Source: (Yahoo Finance, n.d.).



Fig. 5. Closing values of the Russell 2000

Source: (Yahoo Finance, n.d.).

4.1. Finding opportunity windows

As stated above, the first step was to transform the values into categorical data. The data used are mentioned below and presented in Figures 3, 4 and 5. In this analysis, firstly the authors transformed the qualitative data of the index into categorical data using $n = 3, 4, 5$ lengths sliding window-based pattern data. The patterns and their names for different pattern lengths are shown in Table 2.

Table 2. Patterns and their names for different pattern lengths ($n = 3, 4, 5$)

$n = 5$		$n = 4$		$n = 3$	
pattern (0 = down, 1 = up)	pattern name	pattern	pattern name	pattern	pattern name
1	2	3	4	5	6
00000	P1	0000	P1	000	P1
00001	P2	0001	P2	001	P2
00010	P3	0010	P3	010	P3
00011	P4	0011	P4	011	P4
00100	P5	0100	P5	100	P5
00101	P6	0101	P6	101	P6
00110	P7	0110	P7	110	P7
00111	P8	0111	P8	111	P8
01000	P9	1000	P9		

Table 2, cont.

1	2	3	4	5	6
01001	P10	1001	P10		
01010	P11	1010	P11		
01011	P12	1011	P12		
01100	P13	1100	P13		
01101	P14	1101	P14		
01110	P15	1110	P15		
01111	P16	1111	P16		
10000	P17				
10001	P18				
10010	P19				
10011	P20				
10100	P21				
10101	P22				
10110	P23				
10111	P24				
11000	P25				
11001	P26				
11010	P27				
11011	P28				
11100	P29				
11101	P30				
11110	P31				
11111	P32				

Source: authors' calculation.

The analysis started with using the S&P 500 data. Figure 6 indicates the patterns for the S&P 500 index for the three different n . For each different n choice (3, 4, 5) different up and down patterns are generated.

After 2D visualization, it is easier to present the results in a table. Table 3 shows the frequencies for the patterns.

After obtaining the patterns they were compared, Tables 4, 5 and 6 are summarized in this comparison. After the *Down-Up-Up-Up* pattern, the probability of seeing a Down is high (Table 4), while after the *Up-Up-Down-Down* pattern, the probability of seeing an Up is high (Table 4). There is also a statistically significant difference between these proportions using $\alpha = 0.05$ significance level (p -value = 0.023). For Table 5, there are statistically significant differences between the proportions using

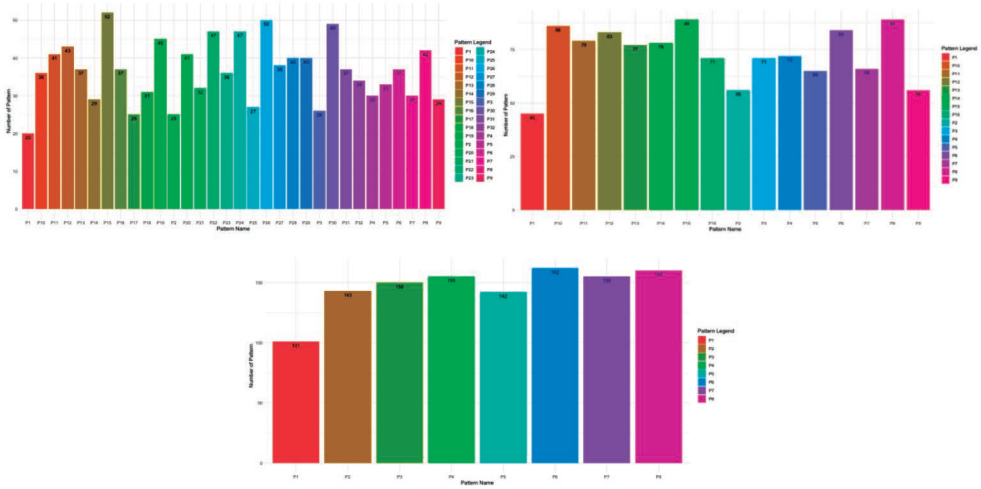


Fig. 6. Number of patterns ($n = 5, 4, 3$) in the S&P 500

Source: authors’ calculation.

Table 3. Patterns and their frequencies for the S&P 500 ($n = 5$)

Pattern (0 = down, 1 = up)	Pattern name	Frequencies
1	2	3
00000	P1	20
00001	P2	25
00010	P3	26
00011	P4	30
00100	P5	33
00101	P6	37
00110	P7	30
00111	P8	42
01000	P9	29
01001	P10	36
01010	P11	41
01011	P12	43
01100	P13	37
01101	P14	29
01110	P15	52
01111	P16	37
10000	P17	25

Table 3, cont.

1	2	3
10001	P18	31
10010	P19	45
10011	P20	41
10100	P21	32
10101	P22	47
10110	P23	36
10111	P24	47
11000	P25	27
11001	P26	50
11010	P27	38
11011	P28	40
11100	P29	40
11101	P30	49
11110	P31	37
11111	P32	34

Source: authors' calculation.

Table 4. The S&P 500 Comparison of Patterns 01110-01111

Pattern	Pattern name	Frequency	Proportions
01110	P15	52	0.58
01111	P16	37	0.42

Source: authors' calculation.

Table 5. The S&P 500 Comparison of Patterns 11000-11001

Pattern	Pattern name	Frequency	Proportions
11000	P25	27	0.35
11001	P26	50	0.65

Source: authors' calculation.

Table 6. The S&P 500 Comparison of Patterns 00110-00111

Pattern	Pattern name	Frequency	Proportions
00110	P7	30	0.42
00111	P8	42	0.58

Source: authors' calculation.

$\alpha = 0.01$ significance level (p -value = 0.000), and for Table 6 there is a statistically significant difference between the proportions using $\alpha = 0.05$ significance level (p -value = 0.043).

Figure 7 shows patterns for the DJIA for $n = 5, 4$ and 3 .

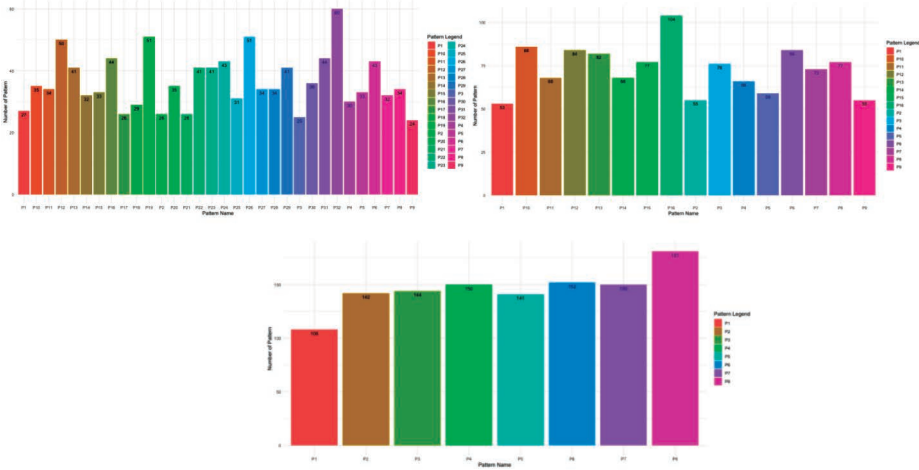


Fig. 7. Number of patterns ($n = 5, 4, 3$) in the DJIA

Source: authors’ calculation.

For the DJIA index, the frequency distribution in Table 7 shows how the frequencies are distributed over the patterns.

Table 7. Patterns and their frequencies for the DJIA ($n = 5$)

Pattern (0 = down, 1 = up)	Pattern name	Frequencies
1	2	3
00000	P1	27
00001	P2	26
00010	P3	25
00011	P4	30
00100	P5	33
00101	P6	43
00110	P7	32
00111	P8	34
01000	P9	24
01001	P10	35
01010	P11	34

Table 7, cont.

1	2	3
01011	P12	50
01100	P13	41
01101	P14	32
01110	P15	33
01111	P16	44
10000	P17	26
10001	P18	29
10010	P19	51
10011	P20	35
10100	P21	26
10101	P22	41
10110	P23	41
10111	P24	43
11000	P25	31
11001	P26	51
11010	P27	34
11011	P28	34
11100	P29	41
11101	P30	36
11110	P31	44
11111	P32	60

Source: authors' calculation.

It was found that the proportions in Table 8 had a statistically significant difference at $\alpha = 0.05$ significance level (p -value = 0.012), in Tables 9 and 10 the proportions also had a statistically significant difference at $\alpha = 0.05$ (p -value = 0.013) and (p -value = 0.025).

Table 8. The DJIA Comparison of Patterns 01010-01011

Pattern	Pattern name	Frequency	Proportions
01010	P11	34	0.40
01011	P12	50	0.60

Source: authors' calculation.

Table 9. The DJIA Comparison of Patterns 10010-10011

Pattern	Pattern name	Frequency	Proportions
10010	P19	51	0.59
10011	P20	35	0.41

Source: authors' calculation.

Table 10. The DJIA Comparison of Patterns 11110-11111

Pattern	Pattern name	Frequency	Proportions
11110	P31	44	0.42
11111	P32	60	0.58

Source: authors' calculation.

Finally, the authors used the Russell 2000 index for the analysis. Figure 8 shows the patterns for the Russell 2000 index.

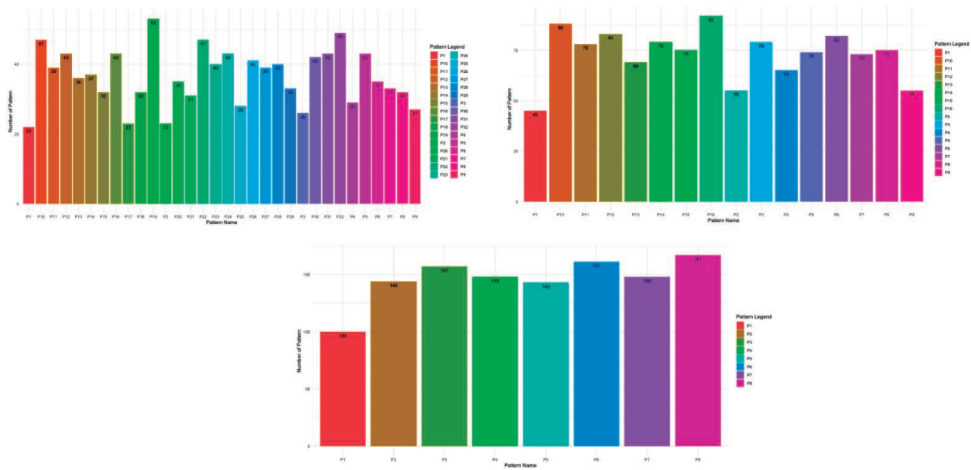


Fig. 8. Number of patterns ($n = 5, 4, 3$) in the Russell 2000

Source: authors' calculation.

Table 11. Patterns and their frequencies for the Russel 2000 ($n = 5$)

Pattern (0 = down, 1 = up)	Pattern name	Frequencies
1	2	3
00000	P1	22
00001	P2	23
00010	P3	26
00011	P4	29
00100	P5	43
00101	P6	35
00110	P7	33
00111	P8	32
01000	P9	27

Table 11, cont.

1	2	3
01001	P10	47
01010	P11	39
01011	P12	43
01100	P13	36
01101	P14	37
01110	P15	32
01111	P16	43
10000	P17	23
10001	P18	32
10010	P19	53
10011	P20	35
10100	P21	31
10101	P22	47
10110	P23	40
10111	P24	43
11000	P25	28
11001	P26	41
11010	P27	39
11011	P28	40
11100	P29	33
11101	P30	42
11110	P31	43
11111	P32	49

Source: authors' calculation.

In Table 12 it was found that the proportions have a statistically significant difference at $\alpha = 0.01$ significance level (p -value = 0.001); in Tables 13 and 14 the proportions also had a statistically significant difference at $\alpha = 0.01$ (p -value = 0.006) and at $\alpha = 0.05$ (p -value = 0.024).

Table 12. The Russell 2000 Comparison of Patterns 01000-01001

Pattern	Pattern name	Frequency	Proportions
01000	P9	27	0.36
01001	P10	47	0.64

Source: authors' calculation.

Table 13. The Russell 2000 Comparison of Patterns 10010-10011

Pattern	Pattern name	Frequency	Probability
10010	P19	53	0.60
10011	P20	35	0.40

Source: authors' calculation.

Table 14. The Russell 2000 Comparison of Patterns 11000-11001

Pattern	Pattern name	Frequency	Probability
11000	P25	28	0.40
11001	P26	41	0.60

Source: authors' calculation.

As seen in Figures 6, 7, and 8, when the pattern lengths were 3 and 4 the variation in pattern frequencies were decreasing, and the small variation in the distribution of pattern frequencies was not enough to create opportunity windows. Therefore, the analyses were performed only for the window length of $n = 5$. Another reason for using $n = 5$ is that this window length corresponds to the number of working days per week.

If the value of n was taken as greater than 5, the pattern numbers increased rapidly as shown in Table 1. In this case, it was also difficult to analyse the pattern frequency distribution tables, and computer codes were also needed for this task. For longer window lengths, these analyses were considered as the subject of another article.

Table 15. What is to be expected on the fifth day

S&P 500		DJIA		Russell 2000	
Pattern	to be expected on the fifth day	pattern	to be expected on the fifth day	pattern	to be expected on the fifth day
0111	decrease	0101	increase	0100	increase
1100	increase	1001	decrease	1001	decrease
0011	increase	1111	increase	1100	increase

Source: authors' calculation.

In the introduction of this paper, the authors stated that only at these points of time the proportions of up and down movements are significantly different from each other, and this can be an opportunity to predict the direction of an index. As seen in Table 15, these points of times for the S&P 500 are the last days of the patterns 0111, 1100, 0011; for the DJIA are the last days of the patterns 0101, 1001, 0011; and finally for the Russell 2000 they are the last days of the patterns 0100, 1001, 1100.

Only three types of patterns can be seen as opportunity windows for the three stock exchange indices. The authors also noted some common patterns for these indices, as 1100 is common for the S&P 500 and the Russell 2000, whilst 1001 is common for the DJIA and the Russell 2000.

5. Conclusion

The aim of this paper was to find the opportunity windows for stock exchange indices. To achieve this the authors chose three stock indices, namely the S&P 500, the DJIA, and the Russell 2000 by applying sliding window techniques used to analyse time series data for the short and long term. One of the most important points in such techniques is the correct determination of the window length. In this study, three different window lengths i.e. 3, 4 and 5 were used, and it was determined that window lengths 3 and 4 were not creating opportunity windows for the prediction of the future direction of the index.

The results show that when the pattern lengths are 3 and 4, creating an opportunity window is difficult to accomplish and it is best to perform the analysis in the 5th length of the pattern. Additionally, when the length was increased after $n = 5$ pattern, the numbers increase rapidly. Finally, the pattern numbers which can be an opportunity to predict the direction of the index, were presented.

References

- Ali, M., Jones, W. M., Xie, X., and Williams, M. (2019). Time cluster: Dimension reduction applied to temporal data for visual analytics. *The Visual Computer*, 35(6-8), 1013-1026.
- Cai, J., and Houge, T. (2008). Long-Term Impact of Russell 2000 Index Rebalancing. *Financial Analysts Journal*, 64(4), 76-91. <https://doi.org/10.2469/faj.v64.n4.7>
- Gareth, J., Witten, D., Hastie, T., and Tibshirani, R. (2017). *An introduction to statistical learning with applications in R*. Springer, Springer Texts in Statistics.
- Hora, S., and Jalbert, T. (2006). The Dow Jones Industrial Average in the twentieth century – Implications for option pricing. *Academy of Accounting and Financial Studies Journal*, 10(3), 17-40.
- Hota, H. S., Handa, R., and Shrivastava, A. K. (2017). Time series data prediction using sliding window based RBF neural network. *International Journal of Computational Intelligence Research*, 13(5), 1145-1156.
- Lerman, D. (2001). *Exchange traded funds and e-mini stock index futures*. Wiley and Sons.
- Makridakis, S., Spiliotis, E., and Assimakopoulos, V. (2018). Statistical and machine learning forecasting methods: Concerns and ways forward. *PLOS ONE*, 13(3), e0194889. doi: 10.1371/journal.pone.0194889
- Mozaffari, L., Mozaffari, A., and Azad, N. (2015). Vehicle speed prediction via a sliding-window time series analysis and an evolutionary least learning machine: A case study on San Francisco urban roads. *Engineering Science and Technology: An International Journal*, 18(2), 150-162. doi: 10.1016/j.jestch.2014.11.002
- Öztürk Katircioğlu, D., Güvenir, H. A., Ravens, U., and Baykal, N. (2017). A window-based time series feature extraction method. *Computers in Biology and Medicine*, (89), 466-486.

- Rajalakshmi, D., and Dinakaran, K. (2015). A survey on effective pattern matching in uncertain time series stream data. *Asian Journal of Applied Sciences*, (8), 217-226. doi: 10.3923/ajaps.2015.217.226
- Senthil, D., and Suseendran, G. (2018). Efficient time series data classification using sliding window technique based improved association rule mining with enhanced support vector machine. *International Journal of Engineering & Technology*, 7(3.3), 218. doi: 10.14419/ijet.v7i2.33.13890
- Siegel, J. J., and Schwartz J. D. (2006). Long-term returns on the original S&P 500 companies. *Financial Analysts Journal*, 62(1) 18-31.
- Spglobal. (2020). Dow Jones Industrial Average®. Retrieved from <https://www.spglobal.com/spdji/en/indices/equity/dow-jones-industrial-average/#overview>
- Sverdllov, A. (2015). *An overview of machine learning and pattern recognition*. Retrieved June 26, 2015 from https://www.gc.cuny.edu/CUNY_GC/media/ComputerScience/Student%20Presentations/Alexander%20Sverdllov/Second_Exam_Survey_Alexander_Sverdllov_6_26_2015.pdf
- Vafaiepour, M., Rahbari, O., Rosen, M., Fazelpour, F., and Ansarirad, P. (2014). Application of sliding window technique for prediction of wind velocity time series. *International Journal of Energy and Environmental Engineering*, 5(2-3). doi: 10.1007/s40095-014-0105-5
- Yahmed Y. B., Azuraliza, A. B., RazakHamdan, A., Almahdi, A., and Abdullah, S. M. S. (2015). Adaptive sliding window algorithm for weather data segmentation. *Journal of Theoretical and Applied Information Technology*, 80(2), 322-333.
- Yahoo Finance (n.d.). Retrieved from https://finance.yahoo.com/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xILmNvbS8&guce_referrer_sig=AQAAAEIeC5nUxiIqbNz7KtFBHz6O9SpJGZNULrSHU51TuFGXN6I2OZ_v6EZkkSgV_7SoQarvGOESNBrlYN2KWsCeqj1tnTebUyflnSY3MwSqUHEXMOWAs9KzWHDVtnpJLqHcy8x77cLPMJc_MQQtq191OAGZp7XT_8_FoxraL8NmmmY
- Yu, Y., Zhu, Y., Li, S., and Wan, D. (2014). Time series outlier detection based on sliding window prediction. *Mathematical Problems in Engineering*, (4). <http://dx.doi.org/10.1155/2014/879736>
- Zhu, Y. and Shasha, D. (2003). Query by Humming: A time series database approach. (Proc. of ACM Special Interest Group on Management of Data). San Diego, California, USA.

WYKORZYSTANIE TECHNIKI SLIDING WINDOW W DANYCH Z SZEREGÓW CZASOWYCH. PRZYKŁAD GIEŁD PAPIERÓW WARTOŚCIOWYCH

Streszczenie: Dane mają swoje formy, a ludzka inteligencja i pojmowanie muszą klasyfikować te formy w celu ich zrozumienia i interpretacji. W niniejszym artykule stosuje się technikę rozsuwanego okna (*sliding window*) i podejmuje próbę odpowiedzi na dwa pytania: czy możliwe jest pojawienie się szansy (*opportunity window*) w szeregach czasowych indeksów giełdowych; jak znaleźć sposób na wykorzystanie pojawiającej się okazji, jeśli taka istnieje. Autorzy zdefiniowali pojęcie *opportunity window* jako okazja (otwarcie) wygenerowana w technice *sliding window*, która może być zastosowana w prognozowaniu. Szukając odpowiedzi, autorzy określili częstotliwości na 3, 4 i 5 długościach wzorców skierowanych w górę i w dół oraz wyjaśnili, jak oszacować okazje osadzone przy użyciu danych szeregów czasowych dla giełd S&P 500, DJIA i Russell 2000. W rezultacie dla S&P 500 ostatnie dni wzorców 0111,1100,0011, dla DJIA ostatnie dni dla 0101, 1001, 0011 oraz dla Russell 2000 ostatnie dni dla 0100, 1001 i 1100 stanowią okazję dla prognozy.

Słowa kluczowe: szeregi czasowe, nauka o danych, wzorce, *sliding window*.