Application of Fuzzy Rough Sets to Financial Time Series Forecasting

Mariusz Podsiadło[1] and Henryk Rybinski[2(✉)]

[1] Misys Plc, London, UK
mariusz.podsiadlo@misys.com
[2] Warsaw University of Technology, Warsaw, Poland
hrb@ii.pw.edu.pl

Abstract. This paper investigates experimentally the feasibility of Fuzzy Rough Sets in building trend prediction models for financial time series, as related research is scarce. Aside of the standard classification accuracy measures, financial profit and loss backtesting using a sample market timing strategy was performed, and profit related quality of the tested methods compared against that of buy&hold strategy applied to the used market indices. The experiments show that Fuzzy Rough Sets models present a viable basis for forecasting market movement direction and thus can support profitable market timing strategies.

Keywords: Artificial intelligence · Rough sets · Fuzzy rough sets · Financial time series prediction

1 Introduction

Forecast of financial markets' conditions is crucial not only for involved institutions and individuals but also for economical well-being of entire nations. On the other hand, financial markets are an inexhaustible source of noisy and incomplete information in a form of financial time series. The number of dimensions, many of which have hidden correlations, and amount of data to be analysed make delivery of good models with sufficient predictive quality very challenging[1]. In this context, a correct forecast of market direction was shown to be sufficient to generate profitable trading strategies [10].

Attempts to employ data discovery and soft computing methods to deliver predictive models have a long track record, including a growing set of reported work based on Rough Sets [13]. Rough Sets and Fuzzy Rough Sets models [14] deliver a way to infer knowledge from noisy and incomplete data and to automatically select significant data features (reducts). This forms a good basis to cope with the challenges of financial time series like data inconsistency and number of dimensions. Rough Sets were shown to be applicable to any time scale, down

---

[1] It is enough to look at the size of quantitative financial engineering teams in any major financial institution.

to intraday trading [9]. However, the recent advances in the field of fuzzy rough sets [14,16] have not been verified in the area of financial time series forecasting yet, although the flexible definition of a similarity relation and ability to work directly with numeric variables make fuzzy rough sets an interesting candidate. Furthermore, related experiments using soft computing models often reported good classification accuracy but financial backtesting and the actual financial performance of the evaluated models were seldom shown. This prompted some reports, where several soft computing models were tested and deemed to deliver performance at best on par with statistical methods and simple buy and hold strategy [7].

Consequently, this work attempts to contribute new research data with regards to the feasibility of rough sets and fuzzy rough sets models to the task of financial time series forecasting. Both, the classification accuracy and financial performance of the examined models were tested using large real life data sets of several well known market-neutral indices. This work extends also on research described in [12] by considering Fuzzy Rough Sets.

This paper is organized as follows: Sect. 2 describes the experiment setup, used data, preprocessing and prediction performance analysis framework. Section 3 presents and discusses the experiment results. Section 4 concludes the paper.

2   Experiment Setup

In this study, Rough Sets (RS), Variable Precision Rough Sets (VPRS) and Fuzzy Rough Sets [14] combined with the k-nearest neighbour algorithm [16] were used to predict direction of movements of US S&P500, German DAX and Hong Kong Hang Seng (HSI) stock market indices[2]. The implementation of the positive region based fuzzy-rough nearest neighbor algorithm [16] provided in [15] was used, with two definitions of the lower/upper set approximation based on:

1. VQRS: Vaguely Quantified Rough Sets [3].
2. Lukasiewicz triangular norm operators, with tolerance relation [5]:

$$R_a(x,y) = 1 - \frac{(|a(x) - a(y)|)}{|a_{max} - a_{min}|} \tag{1}$$

Predictive models based on Variable Precision Rough Sets [17] were used as the reference. A detailed review of roughs sets, fuzzy rough sets, their extensions and their applications in finance is provided in [4,8,13].

All the models were applied to time series data samples of the used stock market indices in order to generate movement predictions using a walk-forward rolling time window cross-validation procedure [6]. Subsequently, using a backtesting procedure applied to Exchange Traded Funds (ETFs) associated with

---

[2] Beating a buy and hold strategy in these efficient markets ought to be challenging.

respective indices, the financial profit and loss of a sample market timing strategy and the *buy and hold* benchmark were compared. The experiment environment was built using R package RoughSets [15], ROSETTA Rough Sets system [11] and an SQL database.

2.1   Input Data

Historical daily Open, Low, High, Close prices, and Volume (OHLC&V) time series of the S&P500, DAX and HSI stock market indices were used to train the models and generate index movement predictions. Additional input was provided by daily closing values of the S&P 500 near-term volatility index VIX[3], which was used to gauge market confidence and covered the below-given periods. Data was divided into training, calibration and test samples, whereas the training sample was set to be 3333 trading days, and the testing period was fixed between 29th of January 2010 and 31st of December 2013 for all indices. This caused a slight difference in the calendar coverage and total sample lengths due to differing holiday calendars of the used indices, i.e. for S&P500 from 1st of August 1996 to 30th of December 2013, DAX from 4th of September 1996 to 27th of December 2013, and HSI from 20th of November 1996 to 30th of December 2013. The associated ETFs, i.e. SPDR S&P500, iShares Core DAX UCITS, and Tracker Fund of Hong Kong, designed to track the performance of the respective indices, were subsequently used to backtest the financial performance of predictions generated for the test sample. The data was adjusted for dividends, splits and mergers. All the used indices and ETF data covered the same time period. Figure 1 displays relative performance of the used index time series over the applied time period.

The input data included 16 conditional variables and one decision variable. The following time series attributes were used as conditional variables: *Open, Close, Volume, VIX Close*. The following derived technical indicators were used as conditional variables: *Acceleration* ($n = 5$), *Average Directional Index* ($n = 50$), *Commodity Channel Index* ($n = 20$), *Chaikin Oscillator* ($n = 3, m = 10$), *Exponential Moving Average* ($n = 50$), *MACD* ($n = 12, m = 26$), *Momentum* ($n = 5$), *Price Oscillator* ($n = 10, m = 5$), *RSI* ($n = 14$), *Price Rate of Change* ($n = 1$), *Williams A/D, Williams %R*. See [1] for detailed description of the indicators[4].

Simple daily return R_i of the given index for a future period i served as the basis for the decision variable d_i (prediction target) defined to be the change direction of the daily simple return R_i as follows:

$$d_i = \begin{cases} down & \text{if } R_i \in (-\infty, 0) \\ up & \text{otherwise} \end{cases} . \tag{2}$$

---

[3] http://www.cboe.com/VIX.
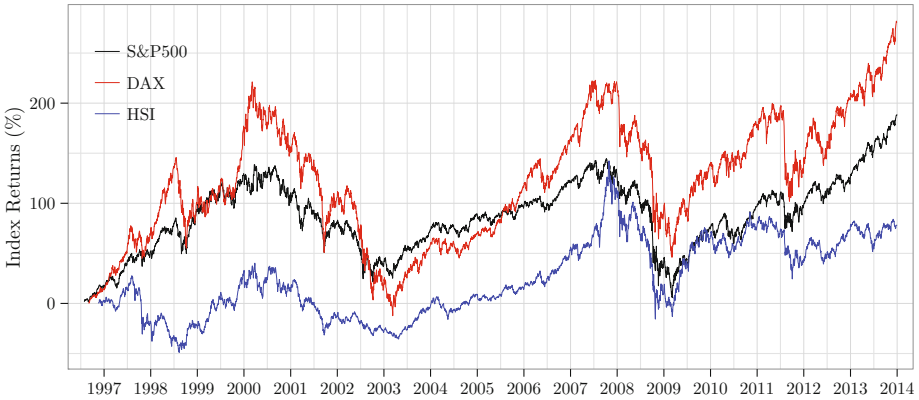[4] The Internet also provides abundance of related information.

Fig. 1. Cumulated return of used indices over the applied time period (Color figure online)

where:

- d_i - the decision variable for i-th sample,
- $R_i = \left(\frac{p_i}{p_{i-1}} - 1\right) * 100$ - simple daily return of the Close price p_i for i-th period (trading day).

The discrete decision variable d_i was used for both, rough sets and fuzzy rough sets classifier models.

2.2   Market Timing Strategy

The financial performance of the tested models was verified using a long-only market timing strategy defined as follows:

$$\text{if } d_i = \begin{cases} down & \text{then } sell \text{ at } Open_{i+1} \\ up & \text{then } buy \text{ at } Open_{i+1} \\ otherwise & \text{then no action} \end{cases} \tag{3}$$

where:

- d_i - decision variable for i-th sample,
- $Open_{i+1}$ - the open price at the next period ($i + 1$-th sample).

Aside of handling the up and down forecast, the strategy accommodated for the case where a prediction model generated no prediction for the forecasted period. In this case the strategy generated no trading signal resulting in no action for the period. The strategy was defined as long only, i.e. no short position was allowed. The *buy and hold* strategy was used as the benchmark strategy. For the purpose of financial performance comparison the ability to buy/sell fractional shares was enabled. This allowed to fully utilize the available equity (initial equity was set to 10'000) in both, buy&sell and market timing strategies. The transaction costs (commissions, slippage, etc.) were ignored in this experiment, as the focus was on the prediction performance.

2.3   Setup and Testing of Classifier Models

All the models were tested with each of selected indices (see Fig. 1) using the walk-forward method with a rolling time window [6]. The walk-forward method derives from the standard cross-validation but also observes the time order of time series and so prevents sample and *look-ahead bias.* For each iteration of the walk-forward cross-validation a consecutively aligned training, calibration and testing samples were selected. The training sample was fixed at 3333 trading days. Multiple combinations of calibration and testing samples were defined drawing their lengths from the sets {21,42,63} and {3,5,21} trading days, respectively.

For each combination of training-calibration-test sample set in the walk-forward cycle, predictive models were created using the training sample. The parameter-driven models, i.e. the fuzzy rough positive region based nearest neighbour model with VQRS approximation ($FPOSNN_{VQRS}$) and the VPRS based model, were then tuned using a grid search on the respective calibration sample, so the optimal (i.e. resulting in the highest classification accuracy/lowest error) model parameter set for the given calibration sample and iteration of the walk forward procedure was found. For the reference VPRS model the tuned parameter was the VPRS $\beta \in \{0.0, 0.25, 0.49\}$. For the $FPOSNN_{VQRS}$ model, the optimal boundaries (α, β) of the *most* quantifier were searched in the set $(\alpha, \beta)_{most} \in \{(0.2, 1), (0.4, 1), (0.6, 1)\}$. The *some* quantifier had fixed boundary parameters set at $(\alpha, \beta)_{some} = (0.1, 0.6)$. The FPOSNN model using the Lukasiewicz $\mathcal{T}$-norm operators ($FPOSNN_{\mathcal{T}_L}$) did not require a specific tuning. The FPOSNN models used the fixed number of neighbours $k = 10$ in the middle of the range suggested in [16].

For VPRS-based analysis the input variables were preprocessed as follows:

1. A subset of technical indicators and the VIX Close values, for which a common consensus on interpretation of their values exists, were discretized using manually defined intervals.
2. The remaining conditional variables were first normalized using the z-score method and subsequently discretized using the equal binning algorithm with 3 bins. The mean and standard deviation of the learning sample as well as its equal binning cuts were used to normalize and discretize the validation and testing samples.

For the fuzzy rough sets-based processing all conditional variables were normalized, like in case of the rough sets model. Fuzzy rough sets do not require discretization of conditional variables. The decision variable d (Eq. 2) was encoded as 1 for an *up* movement, and -1 for a *down* movement. The VPRS model set the decision variable d to 0 in case, the ruleset was unable to predict its value (i.e. no matching rule was found, as per Eq. 3).

3   Experiment Results

The walk-forward procedure was executed using rough sets and fuzzy rough sets models generated for all combinations of test sample lengths and model specific

Table 1. Prediction performance of tested models

Test Sample Length	Test Sample Hit Ratio (%)								
	$VPRS$			$FPOSNN_{VQRS}$			$FPOSNN_{T_L}$		
	S&P500	HSI	DAX	S&P500	HSI	DAX	S&P500	HSI	DAX
Calibration period = 21									
3	53.09	**50.51**	51.64	*54.20*	50.10	51.44	53.90	50.20	50.65
5	52.29	**50.51**	**51.74**	53.60	50.00	51.54	53.19	49.80	50.94
21	52.58	**50.51**	49.95	53.90	49.70	51.64	52.89	50.10	51.44
Calibration period = 42									
3	51.87	50.61	49.85	52.89	50.81	52.14	53.18	51.11	52.04
5	51.25	50.71	50.55	**53.09**	50.61	52.04	51.87	51.42	**52.23**
21	50.30	**51.62**	48.96	51.17	50.20	51.84	50.25	50.30	52.14
Calibration period = 63									
3	50.56	52.13	49.35	**52.89**	50.00	52.83	52.28	50.10	52.23
5	51.47	51.52	49.85	52.28	50.30	52.83	52.08	50.71	52.53
21	49.85	*52.43*	49.85	52.48	49.80	*52.93*	52.58	50.10	52.63

parameters as defined in Sect. 2.3. Both, the predictive and financial performance were tested with results described below.

3.1   Prediction Performance

Table 1 presents the total classification performance of the generated models. The prediction models based on $FPOSNN_{VQRS}$ delivered classification performance with the highest accuracy for S&P500 and DAX indices. i.e. 54.21 % (calibration sample/forecast period=21/3 trading days) and 52.93 % (calibration sample/forecast period=63/21 trading days), respectively. The VPRS rule based model delivered the best accuracy for HSI with 52.43 % vs. the worst performing $FPOSNN_{VQRS}$ model with classification accuracy of 49.80 % for the calibration sample/forecast period=63/21 trading days. The FPOSNN models consistently outperformed the VPRS rule based model for all tested combinations of calibration and forecast periods for the S&P index. For S&P and DAX time series the $FPOSNN_{VQRS}$ model delivered the best average accuracy of 53.11 % and 52.14 %, respectively, vs. the respective average performance of the VPRS rule based model of 51.54 % and 50.90 %. For the HSI index the $FPOSNN_{VQRS}$ and $FPOSNN_{T_L}$ models delivered an average classification performance of 50.17 % and 50.43, vs. 51.17 % delivered by the VPRS model. The length of calibration and forecast periods had an impact on the classification efficiency for all tested models. The use and increasing length of the calibration sample had a positive impact only on performance of the VPRS model, when predicting movements of the HSI index. For other indices the increasing length of the calibration sample had none to negative impact on classification accuracy of the tested models. This shows that the time distance between the classified sample and the ruleset or the nearest neighbour base sample (i.e. the training sample) was more important than the grid search based tuning of model parameters. The latter required a calibration sample, thus increasing the time distance between the training and

test samples. One can conclude that consideration of time distance as a factor in the classification process would most likely improve it [12]. It is also worth noticing that both, DAX and S&P500 had a strong upwards trend in more than a half of the testing period, whereas HSI growth was relatively moderate with a sidewards trend in the same time period (see Fig. 2). Thus, DAX and S&P500 time series supported trend following prediction models, and by extension the nearest neighbour algorithm taking decisions based on a relatively small number of k samples. Prediction of sidewards movements (so called *whipsaws*), like it was the case for the HSI time series, is recognised as challenging in the field of technical analysis. In this case the rule based algorithm was relatively better. The classification accuracy alone is however not sufficient to deliver a profitable predictive model. A decisive factor is the *timing* of the correct prediction, which directly impacts the financial performance.

3.2   Financial Performance

The profit and loss backtesting using the market timing strategy outlined in Sect. 2.2 as well as the *buy and hold* benchmark strategy were executed for all
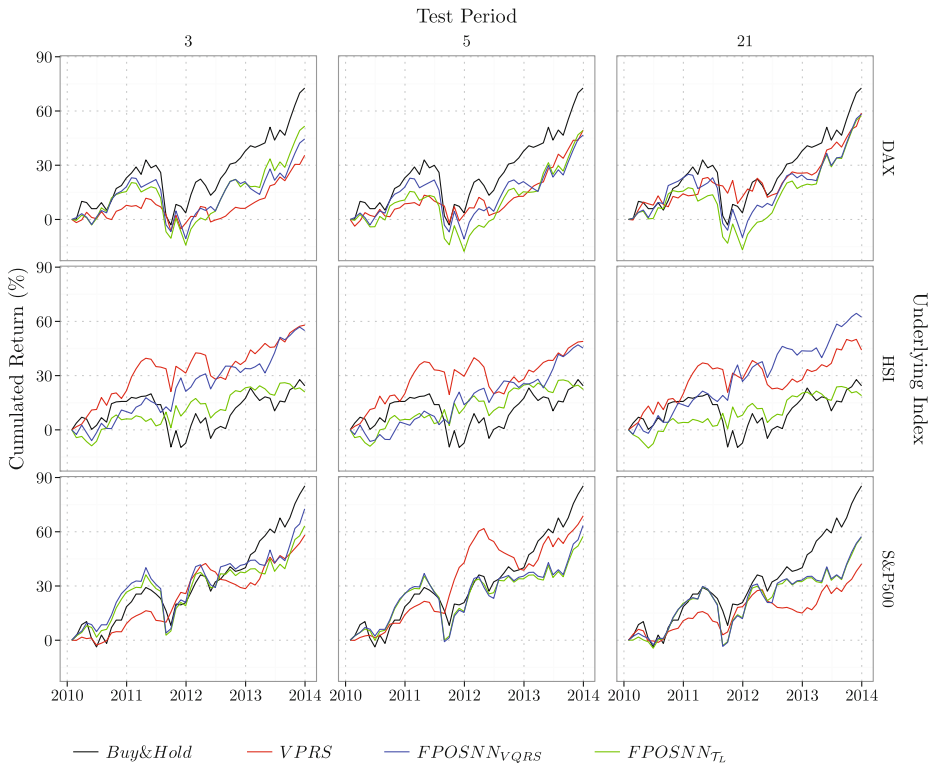

Fig. 2. Cumulated return time series (calibration sample of 63 trading days) (Color figure online)

combinations of the training, calibration and test sample lengths. The trading simulation used tracking ETFs associated with the underlying indices as the traded equity instrument. The trading costs and slippage were ignored. Figure 2 displays the cumulated return time series of tested models vs. buy and hold strategy for test runs using the calibration period of 63 trading days[5], whereas Table 2 displays related financial performance summary. The S&P500 and DAX indices were in a strong upwards trend for the second half of the testing period, reflected in the total return of 85.39 % and 72.69 % of the respective tracking ETFs (and associated buy and hold strategy). Consequently, beating these indices was expected to be difficult. This was confirmed for the majority of the calibration/test sample length combinations, where the predictive models underperformed the market, although still generating positive returns above 30 % (DAX) and 40 % (S&P500). An exception was the $FPOSNN_{VQRS}$ model, consistently beating the S&P500 performance when using the calibration period od 21 days, and delivering the peak return of 99.88 % using the 21/21 combination of calibration/test sample lengths. A similar exceptional performance in case of the DAX benchmark was delivered by the VPRS model with 90.13 % and 78.68 % return using the 42/3 and 42/5 trading days combination of calibration/test, respectively. In case of the S&P500 index, the decrease of financial return was coupled with the increased length of the calibration period, reflecting the similar observation done with regards to the classification accuracy. For the DAX index, this trend was not present, with the VPRS model performing the best on the calibration sample of 42 days. It should be noted, that DAX movements were more volatile than these of S&P500, although both indices shared the same strong upward trend in the second half of the testing period. This may explain the reason for the lack of the clear dependencies visible in the case of the S&P500 index, i.e. the US VIX volatility index might not be able to fully reflect the volatility of the German DAX index. Thus, it may be necessary to define more localized input variables for the tested models, so the inherent capability of Rough Sets to select reducts is applied for each underlying market. The HSI index reflects the above observations amplified by the fact that it had a relatively weaker upwards trend, with multiple whipsaws, resulting in the total return of the tracking ETF of only 24.36 % over the testing period. Consequently, the correct prediction of the index movement played larger role and can be seen in an excellent performance of $FPOSNN_{VQRS}$ and VPRS models outperforming the Hong Kong market for all combinations of tested lengths of calibration and test samples. Both underlying models, namely VPRS and VQRS are closely related, aiming at reduction of classification noise caused by outliers. On the other hand, the $FPOSNN_{\mathcal{T}_L}$ model underperformed the market, which would indicate that this model is actually trend following and so unable to work in a nontrending market. Even though the tested models mostly underperform DAX and S&P500 markets due to the above described dependencies, the total return alone gives a biased view of the performance without considering the associated risk.

---

[5] The full data set was not shown due to space limitations but can be obtained from the authors.

Table 2. Financial performance of tested models on a holdout sample

Model	Cum. Return			Information Ratio			Max. Drawdown		
	HSI	DAX	S&P500	HSI	DAX	S&P500	HSI	DAX	S&P500
Calibration period = 63, Test period = 3									
$VPRS$	**57.96**	35.50	58.16	*0.46*	-0.43	-0.37	18.18	**23.01**	*12.17*
$FPOSNN_{VQRS}$	54.92	44.54	**72.68**	0.21	-0.19	*-0.11*	**14.76**	32.03	33.66
$FPOSNN_{T_L}$	20.92	**51.62**	63.25	−0.07	**-0.14**	-0.13	15.90	33.27	32.54
Calibration period = 63, Test period = 5									
$VPRS$	**48.94**	**49.34**	**68.83**	**0.34**	-0.26	-0.22	18.18	*20.60*	**16.50**
$FPOSNN_{VQRS}$	45.22	46.60	63.47	0.15	-0.18	**-0.17**	14.76	32.03	35.24
$FPOSNN_{T_L}$	22.40	49.19	57.45	−0.04	**-0.16**	-0.18	**14.64**	33.27	34.02
Calibration period = 63, Test period = 21									
$VPRS$	44.14	**58.88**	42.17	**0.28**	-0.15	-0.60	16.90	**20.82**	**14.96**
$FPOSNN_{VQRS}$	*62.42*	58.25	**57.34**	0.25	*-0.09*	-0.22	*13.21*	32.45	33.47
$FPOSNN_{T_L}$	18.92	57.37	56.87	−0.10	-0.10	**-0.18**	14.87	33.68	33.07
Buy&Hold	24.36	*72.69*	*85.39*	-	-	-	32.99	32.65	18.61

Especially, the information ratio and maximal drawdowns [2] showed the level of risk associated with the permanent exposure required by the buy&hold strategy. The FPOSNN based models had risk similar to that of DAX and S&P but performed well vs. the HSI index, with less than 50 % of the drawdown caused by the index. The VPRS rule based model also delivered performance slightly lower than DAX and S&P500 indices but with a much lower level of risk (drawdown). Like the $FPOSNN_{VQRS}$ model, it outperformed the HSI index with the half of the risk, reflected in the 50 % lower drawdown and positive information ratio.

4   Conclusions

Variable Precision Rough Sets and Fuzzy Rough Sets were used to generate binary classifiers applied to real life market data of multiple stock indices. Both, the classification accuracy and ability to support profitable market timing strategies were evaluated. The experimental results showed that Fuzzy Rough Sets based models were able to outperform the VPRS based model in terms of classification accuracy in the majority of experiments. In terms of financial performance, the VPRS and $FPOSNN_{VQRS}$ models were able to outperform the buy&hold strategy applied to the HSI index. Considering the simplicity of the used strategy and the strong upwards trend of S&P500 and DAX, the models were robust when applied to these times series. The application of Rough Sets models to portfolios representing a specific strategy can be simulated by tracking certain indices or exchange traded funds and is planned to be included in the future research. In general, the effectiveness of the delivered trading signals can be further improved by considering the time distance in classification algorithms and wider set of input variables.

References

1. Achelis, S.: Technical Analysis from A to Z, 2nd edn. McGraw-Hill Education, New York (2000)
2. Bacon, C.R.: Practical Risk-Adjusted Performance Measurement. Wiley, Oxford (2012)
3. Cornelis, C., De Cock, M., Radzikowska, A.M.: Vaguely quantified rough sets. In: An, A., Stefanowski, J., Ramanna, S., Butz, C.J., Pedrycz, W., Wang, G. (eds.) RSFDGrC 2007. LNCS (LNAI), vol. 4482, pp. 87–94. Springer, Heidelberg (2007)
4. Dubois, D., Prade, H.: Rough fuzzy sets and fuzzy rough sets. Int. J. Gen. Syst. **17**(2–3), 191–209 (1990)
5. Jensen, R., Cornelis, C., Shen, Q.: Hybrid fuzzy-rough rule induction and feature selection. In: 2009 IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2009, pp. 1151–1156. IEEE (2009)
6. Kaastra, I., Boyd, M.: Designing a neural network for forecasting financial and economic time series. Neurocomputing **10**(3), 215–236 (1996)
7. Kinlay, J., Rico, D.: Can machine learning techniques be used to predict market direction?-the 1,000,000 model test (2011)
8. Komorowski, J., Pawlak, Z., Polkowski, L.: Rough sets: a tutorial. In: Pal, S.K., Skowron, A. (eds.) Rough Fuzzy Hybridization: A New Trend In Decision-Making, pp. 3–98. Springer, Singapore (1999)
9. Lee, S.J., et al.: Using rough set to support investment strategies of real-time trading in futures market. Appl. Intell. **32**(3), 364–377 (2010)
10. Leung, M.T., Daouk, H., Chen, A.S.: Forecasting stock indices: a comparison of classification and level estimation models. Int. J. Forecast. **16**(2), 173–190 (2000)
11. Øhrn, A.: Rosetta technical reference manual, Department of Computer and Information Science, pp. 1–66. Norwegian University of Science and Technology (NTNU), Trondheim (2000)
12. Podsiadło, M., Rybiński, H.: Financial time series forecasting using rough sets with time-weighted rule voting. Rep. of Inst. of Comp. Sci., 1/2015, Warsaw University of Technology, subm. to. Eur. J. Oper. Res
13. Podsiadło, M., Rybiński, H.: Rough sets in economy and finance. In: Peters, J.F., Skowron, A. (eds.) Transactions on Rough Sets XVII. LNCS, vol. 8375, pp. 109–173. Springer, Heidelberg (2014)
14. Radzikowska, A.M., Kerre, E.E.: A comparative study of fuzzy rough sets. Fuzzy sets syst. **126**(2), 137–155 (2002)
15. Riza, L.S., et al.: Implementing algorithms of rough set theory and fuzzy rough set theory in the R package RoughSets. Inf. Sci. **287**, 68–89 (2014)
16. Verbiest, N., Cornelis, C., Jensen, R.: Fuzzy rough positive region based nearest neighbour classification. In: 2012 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1–7. IEEE (2012)
17. Ziarko, W.: Variable precision rough set model. J. comput. syst. sci. **46**(1), 39–59 (1993)