Journal of Economics and Administrative Sciences (JEAS)

Available online at http://jeasiq.uobaghdad.edu.iq

Estimating Stock Returns Using Rough Set Theory: An Exploratory study With An Evidence From Iraq Stock Exchange

Assistant lecturer: Mohammed H. Adnan
Northern Technical University , Mosul, Iraq
mohammed.hamdan@ntu.edu.iq

Prof. Mustafa Muneer Isma'eel, PhD
University Of Baghdad / College Of Administration & Economics , Baghdad, Iraq
Mustafa.m@coadec.uobaghdad.edu.iq

Received:28/4/2021 Accepted: 30/5/2021 Published: June/ 2021

This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International (CC BY-NC 4.0)

ABSTRACT

This research aims to estimate stock returns, according to the Rough Set Theory approach, test its effectiveness and accuracy in predicting stock returns and their potential in the field of financial markets, and rationalize investor decisions. The research sample is totaling (10) companies traded at Iraq Stock Exchange. The results showed a remarkable Rough Set Theory application in data reduction, contributing to the rationalization of investment decisions. The most prominent conclusions are the capability of rough set theory in dealing with financial data and applying it for forecasting stock returns.The research provides those interested in investing stocks in financial markets with significant financial analysis tools that exceed the traditional statistical methods. The originality of the research lies in the diversification of financial and statistical analysis tools and methods of forecasting stock returns.

Keywords: Stock Returns, Rough Set Theory, Conditional Attributes, Decision Rule, lower approximation, upper approximation..

* Research extracted from the doctoral thesis of the first researcher

1. Introduction

The activities of common stock trading are based on expectations of their future returns, which are reflected in the importance of estimating the future returns of stock to enable investors to make investment decisions in common stocks to achieve the highest possible returns resulting from those investments. Investors are increasingly trying in various classical and modern financial theorie s and in the presence of financial markets to find effective and efficient methods for estimating future returns of a stock.

As investors resort to technical analysis methods of the financial markets concerned with estimating stock indices and their movement using historical data and representing them in graphical forms, as well as basic economic analysis that deals with prominent economic indicators in the financial market.

Mathematical models and equations were also numerous to estimate the return of stock required by investors, such as the Capital Asset Pricing Model (CAPM), and other models were used to estimate future returns on stock that contribute to making investment decisions.

Several recent studies have emerged to use more diverse methods for estimating future returns of stock, including the use of genetic algorithms and fuzzy sets, as well as rough set theory and neural networks in an attempt to obtain more accurate results in estimating future returns of a stock. Which called for a compelling study of the rough set theory in estimating the expected returns of stocks.

The study (Shen & Loh, 2004) indicated applying the Rough Set Theory for forecasting stock indices and that it is effective for achieving prediction.

It was evident from the study (Shyng et al., 2011) that the application of Rough Set Theory reduces the amount of data and provides alternative strategies that can help decision-makers in analyzing the data.

For analyzing business indicators using Rough Set Theory, the study (Couto, 2015) indicated that it provides a more critical view of investment decisions and measuring the performance of potential competitors. To determine futures trading strategies, the study (Kim & Enke, 2016), in applying the theory of Rough Set, confirmed the achievement of profitable results for both the buy and hold strategy.

In a study (Cheng & Yang, 2018) to predict stock price using a fuzzy time series model based on the extrapolation of an approximate set rule, the proposed method outperformed the list models in terms of error indicators and profits.

In order to predict the time series of stock market forecasts, the study (Pal & Kar, 2019) came through the fuzzy estimation data and the generation of rules using Rough Set Theory and confirmed that the method is more effective by using the Rough Set Theory in generating rules Decision.

Thus, most studies have indicated the effectiveness of the Rough Set Theory in predicting stock prices and various business and finance fields. On this basis, interest in the topic of research was preferred in an attempt to apply scientific fields, models, methods, and a variety that support the field of business administration, especially the financial field, because of its great importance in the field of business and its nature affected by multiple fields of knowledge.

The choice was to use rough set theory in estimating the future returns of stocks for a sample of the Iraq Stock Exchange companies and to ensure their accuracy in estimating the results.

The cognitive structure of research supports financial thought and improves forecasting processes due to its cognitive properties. Thus, the research seeks to know, apply and explore modern models and methods in financial markets in an attempt to achieve a cognitive advantage in this field as well as support researchers and those interested in the field—financial markets with those models.

Based on that, the importance of research is highlighted to help those interested in financial thought and specialists in financial markets in diversifying methods of forecasting future returns of stocks.

In order to achieve the objectives of the research and study its idea at the theoretical and experimental level, the research in the first section included an introduction and a set of previous studies in financial thought that included the use of rough set theory .

The second section includes the theoretical aspect of the research, the framework of Rough Set Theory, and the rationale for its use in estimating inventory returns and measuring tools.

The third section included a presentation of the research methodology as well as research data and experiment method.

The fourth section deals with the experimental aspect of the research represented in the results of estimating returns using rough set theory , as well as a discussion of the results of the research.

The fifth section contains conclusions that convert the results of quantitative measurement into intellectual content that can be inferred.

2. Rough Set Theory and justifications for its use to Estimate stock returns

Pawlak (1982) made one of his contributions to mathematics and quantitative methods embodied by the Rough Set Theory, which was employed for forecasting, including forecasting stock prices, their returns, and discovering trading signals and rules (Cheng et al., 2010). The theory helps in decision-making in the conditions of uncertainty that characterize the financial markets and stock investments. The Rough set approach assumes that each element has a certain amount of related information expressed through the features describing stocks and securities in general and expressed in the elements of the set (Liao & Chang, 2016).

Pawlak (2002) described Rough Set Theory as a new mathematical tool for analyzing imperfect data, and the theory has been applied in many fields of decision support, engineering, the environment, banking, financial forecasting, medicine, and others (Pawlak, 2002).

Rough Set Theory has contributed to various types of inferential thinking styles, including their use for financial forecasting purposes (Pawlak & Skowron, 2007). This theory has its advantages that can be summarized as follows (Podsiadło, 2016; Tay & Shen, 2002; Kalaivani et al., 2017):

1) Provides mathematical methods to discover patterns of ambiguous data.
2) The ability to propose an ideal set of data, metadata, and data reduction.

3) Availability of quantitative evaluation of the importance of data relative to the specified attributes.
4) The ability to derive decision rules from the used dataset.
5) Adoption of accurate data.
6) It provides a set of decision rules derived from the Rough Set model describing the financial information.
7) The decision rules obtained from the Rough Set model are based on facts.
8) The results of the Rough Set model are easy to understand.

These advantages make Rough set an attractive model of important forecasting models for financial chains, including stock return forecasts, as influences, large data volumes, and frequency of change are inherent in them. Moreover, the Rough Set model deals with primary data as it is a private community and does not require any additional information or a combination thereof (Podsiadło, 2016).

2.1. <u>The decision table is a framework for Rough Set theory and the nondiscrimination relationship</u>

Pawlak developed decision logic to derive knowledge related to accurate data by the method of representing knowledge as a table that includes a set of attributes to distinguish and classify the data, called a knowledge representation system which has the advantages of representing knowledge in the form of a table. The spreadsheet can be interpreted differently; that is, it can be formalized. It is a logical system for decision-making (Akama et al., 2018). Each row within the information system or the decision table represents an element, while each column embodies a measurable characteristic of the same element (Upadhyaya et al., 2006).

The information system $I = (U, C)$ may contain extra information, and its reduction does not affect the objectivity of the classification. The Rough set theory determines the classification of categories using the concept of the non-discrimination relationship according to the following equation (Podsiadło & Rybinski, 2016):

$$IND_B = \{(x, y) \in U^2, B \subseteq A \,|\, \forall_a \in B\ a(x) = a(y)\} \quad \ldots(1)$$

The INDB relationship is an equivalence relationship, meaning that it is reflexive, symmetric, and multiple; as can be seen from the previous equation, if $(x, y) \in IND_B$, then the elements (x, y) Belong to the same department $U(U/IND_B \text{ or}, U/B)$. Meaning that (x, y) are indistinguishable from each other concerning the traits of B, and they belong to the same class of the set of elementary traits B, denoted as $[x]_B$, $x \in U$, where [x] represents the equivalent rows of x.

2.2. <u>Measurements of Rough set theory and its processes</u>

The approximation process represents a method of classifying the elements of the data community according to the specific characteristics and the characteristics and values for each element with possible explanations and types of Rough set of elements (Pagliani & Chakraborty, 2008).

The approximate space forms the basic idea of Rough set theory expressed by the approximation distances for the decision table data or the information system (Hassanien et al., 2007). The approximate space of a set is created from the classification of a subset instead of a set of the elements of the entire set so

that the search for the elements of the approximate space related to them becomes (Skowron et al., 2012).

Based on the determination of the indistinguishable relationship and the equivalence class of the elements, the elements of a whole society are calculated and classified according to the following concepts and measurements (Chen & Tsai, 2016; Suraj, 2004; Li, 2014):

1) lower approximation: $\underline{B}(X) = \{x \in U: [x] \subseteq X$ is called approximation of the lower resolution X . According to the attribute B. Or the positive region of X. Represents a clear set of elements belonging to the data population U and can be classified as belonging to the decision rule X using the characteristics set of attribute B.

2) Upper approximation: $\overline{B}\,(X) = \{x \in U: [x]_B \cap X \neq \emptyset\}$. the upper approximation of resolution X is called the attribute B and includes a clear set of the sum of the elements within the data population U, which can be classified as belonging to X using the character set B and whose intersection with the resolution X represents a non-null set.

3) Boundary region : $BND_B(X) = \overline{B}(X) - \underline{B}(X)$ It is called the boundary region of X . it is the set of elements within the data community U that cannot be categorized with certainty as belonging to or not belonging to X using the set of attributes B. The uncertainty relates to the classification of data within the boundary region, which means the presence of elements that cannot Classify them uniquely concerning a set or its complement (Chikalov . et al.,2012).

4) Negative region: The negative region B is a set of elements within the data population U, which certainly cannot be classified as belonging to X using the attributes set B, set X with $BND_X \neq \emptyset$.

5) Accurate measure: The accuracy measure represents the degree of completeness of knowledge about the set of elements specified in the decision class X and is determined by the ratio of the sum of the elements of the essential elements to the lower and upper approximations of X as follows (Ma & Mi,2016 ; Cao et al., 2003):

$$a_{B(x)} = \frac{|\underline{B}(X)|}{|\overline{B}(X)|} \quad \ldots\ldots(2)$$

$$a_{B(x)} < 1$$

Figure (1) shows the limits of the Rough set.

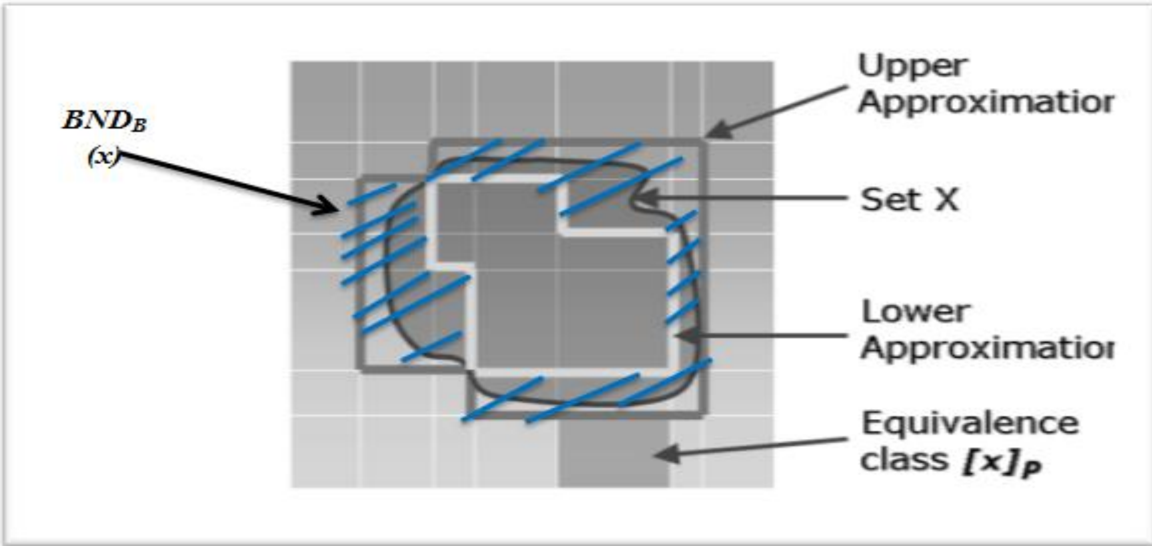

Figure (1) Rough set boundaries

Source:
Shen, Qiang Jensen, Richard (2007)" Rough Sets, their Extensions and Applications," International Journal of Automation and Computing, P(2).

3. Methodology

The experimental aspect of the research includes the pillars of applying quantitative and experimental analysis of the research. It appeared from the research sample, the nature of the data, the time horizon, and the results of forecasting stock returns using rough set theory , and discussing them in comparison with previous studies.

3.1. Research sample and data

The research data included indicators (highest price, lowest price, closing price) monthly for stocks from historical data bulletins for the Iraq Stock Exchange for the research sample consisting of (10) stocks of industrial sector companies listed on the Iraq Stock Exchange that were selected according to the stocks of relevant companies. The most minor interruptions in trading for the period from January 2010 to December 2019 occurred (120) months, and the search period of (120) months included interruptions in trading in the stocks of companies. The research sample led to the loss of some values of the indicators adopted in the research. The stocks result from monthly trading interruptions, and the missing value data for stock prices have been compensated by the double exponential smoothing method using the (Minitab17) program.

3.2. Executing the process of estimating stock returns

The Rough set theory was applied to the research sample data using the (MATLAB 12) program and building an algorithm for data classification based on rough set theory. They used conditional attributes to classify the data into two main categories expressed by high or low prices for each month over the previous month. For the approved indicators for forming the Rough set, calculating the lower approximation, upper approximation, and boundary region. As well as measures of approximation accuracy and the degree of dependence of attributes

in the classification of data, while inconsistent sets were excluded from the decision table from the category of the approved set for forecasting after classifying set. Forecasting the returns of stocks After calling the closing prices corresponding to the lower approximate set and calculating the average rate of return for the change in the stock price for the less approximate set elements to estimate the stock returns. The return on stocks of the research sample companies was calculated according to the following formula (Qiu et al., 2016):

$$y_t = \frac{P_t - P_{t-1}}{P_{t-1}} \quad .....(3)$$

4. Results and discussions

Table (1) displays the stock returns for companies in the industrial sector and from it the decision categories (A, B), which include the number of elements (months), the least approximation for the two categories, measures of approximation accuracy, and the degree of dependency. The results indicate the reliability of the decision base (A) for predicting stock returns.

As it possible, measures of approximation accuracy better than measures of approximation accuracy for rule (B). Which indicates an inconsistency in the results of approximation. This calls for excluding decision rule (B) from the elements of the standard set for prediction as it collectively meets the criterion Accuracy greater than (1) the correct one, which means that it is excluded from the prediction set because the number of approximation elements is greater than the initial set of approximation. While the results of the decision rule (A) indicate, the consistency of the approximation results and the rule (A) can be adopted.

For forecasting stock returns, the accuracy of the approximation criterion is less than (1) correct, Which means fulfilling the condition that the elements (months) of the lowest approximate group of the decision rule (A) are part of the total data set and that they have the same conditional attributes for classifying the elements and not vice versa.

The elements (months) of the approximate set were approved from the decision rule (A) for calculating and estimating the rate of return for stocks, the research sample, and excluding the category (B), and they are clarified as follows:

No	Stock name	Return	Rule A			Rule B		
			Lower	Accuracy	Depend	Lower	Accuracy	Depend
1	Al-Mansour Pharmaceuticals Industries	-0.035	35	0.50	0.58	85	1.70	0.42
2	Modern Sewing	0.101	45	0.56	0.65	75	1.83	0.34
3	Iraqi For Tufted Carpets	0.035	23	0.32	0.58	97	1.94	0.42
4	Baghdad for Packing Materials	-0.025	34	0.50	0.55	86	1.62	0.44
5	Baghdad Soft Drinks	0.066	35	0.47	0.61	85	1.85	0.38
6	National Chemical &Plastic Industries	-0.019	38	0.51	0.61	82	1.78	0.38
7	AL- Kindi of Veterinary Vaccines	0.016	40	0.52	0.63	80	1.82	0.37
8	Iraqi Engineering Works	-0.019	41	0.51	0.66	79	1.98	0.33
9	Metallic Industries and Bicycles	0.015	51	0.63	0.66	69	1.73	0.33
10	Ready Made Clothes	0.075	31	0.45	0.56	89	1.71	0.43

The industry sector within the research sample contained (10) stocks whose data were analyzed according to the approximate set theory. The future returns of their stocks were estimated, respectively, to reach (-0.035 ,0.101, 0.035 ,-0.025 ,0.066 ,-0.019,0.016,-0.019,0.015,0.075), which was calculated according to the closing prices of the set's approximate elements that are the lowest for the classification of the data of each stock by the numbers of the elements, respectively (35,45,23,34, 35,38,40,41,51,31).

It was clear from the results of the rounding accuracy measurement to classify the same stocks, respectively (0.50, 0.56, 0.32, 0.50, 0.47, 0.51, 0.52, 0.51, 0.63 , 0.45) . In an expression of the ratio of the lowest approximation to the highest approximation and the extent of data reduction between the two approximations, while the degree of dependence of the attributes on the same stocks respectively (0.58, 0.65, 0.58, 0.55, 0.61, 0.61, 0.63, 0.66, 0.66, 0.56). In an expression of the ability of conditional attributes to classify data according to Rough set theory.

The research results are consistent with a study (Yu Shyng et al., 2011) that the application of Rough Set Theory reduces the amount of data and provides alternative strategies that can help decision-makers analyze the data.

For analyzing business indicators using Rough Set Theory, the study (Couto, 2015) indicated that it provides a more important view of investment decisions and measuring the performance of potential competitors.

5. Conclusions

The rough set theory provides great possibilities in classifying data according to the conditional attributes that classify data based on the relationship of non-discrimination. It helps in the reduction of data which saves excellent effort and time in obtaining sufficient information to make decisions by investors in financial markets, a capability the rough set theory in dealing with financial data and applying it to forecast stock returns—supporting those who are interested in Financial thought using modern mathematical methods that can be used in the field of financial markets and to forecast stock returns. The accuracy of the rounding results depends on the preparation of the conditional attributes, which are the basis of the classification process.

Reference

1.Akama, S., Murai, T., & Kudo, Y. (2018). Reasoning with Rough Sets. Logical Approaches to Granularity-Based Framework.

2.Cao, L., Shen, L., & Tay, E. H. F. (2003). Ordinary shares, exotic methods: Financial forecasting using data mining techniques. World Scientific.

3.Cheng, C. H., & Yang, J. H. (2018). Fuzzy time-series model based on rough set rule induction for forecasting stock price. Neurocomputing, 302, 33-45.

4.Cheng, C. H., Chen, T. L., & Wei, L. Y. (2010). A hybrid model based on rough sets theory and genetic algorithms for stock price forecasting. Information Sciences, 180(9), 1610-1629.

5.Chikalov, I., Lozin, V., Lozina, I., Moshkov, M., Nguyen, H. S., Skowron, A., & Zielosko, B. (2012). Three approaches to data analysis: Test theory, rough sets and logical analysis of data (Vol. 41). Springer Science & Business Media.

6.Couto, A. B. G. (2015). Using a Dominance-based Rough Set Approach for Analysing Business Indicators. In ITQM (pp. 350-359).

7.Hassanien, A. E., Suraj, Z., Slezak, D., & Lingras, P. (Eds.). (2007). Rough Computing: Theories, Technologies and Applications: Theories, Technologies and Applications. IGI Global.

8.Kalaivani.R M.V.Suresh & N.Srinivasan(2017) . A Study of Rough Sets Theory and It's Application Over Various Fields. Journal of Applied Science and Engineering Methodologies, 3 (2).

9.Kim, Y., & Enke, D. (2016). Developing a rule change trading system for the futures market using rough set analysis. Expert Systems with Applications, 59, 165-173.

10. Li, X. (2014). Attribute selection methods in rough set theory.

11. Liao, S. H., & Chang, H. K. (2016). A rough set-based association rule approach for a recommendation system for online consumers. Information Processing & Management, 52(6), 1142-1160.

12. Ma, Z., & Mi, J. (2016). Boundary region-based rough sets and uncertainty measures in the approximation space. Information Sciences, 370, 239-255.

13. Pagliani, P., & Chakraborty, M. (2008). A geometry of approximation. Springer.

14. Pal, S. S., & Kar, S. (2019). Time series forecasting for stock market prediction through data discretization by fuzzistics and rule generation by rough set theory. Mathematics and Computers in Simulation, 162, 18-30.

15. Pawlak, Z. (2002). Rough set theory and its applications. Journal of Telecommunications and information technology, 7-10.
16. Pawlak, Z., & Skowron, A. (2007). Rudiments of rough sets. Information sciences, 177(1), 3-27.
17. Podsiadło, M. (2016). Forecasting Financial Time Series Movements with Rough Sets and Fuzzy Rough Sets (Doctoral dissertation, The Institute of Computer Science).
18. Podsiadlo, M., & Rybinski, H. (2016). Financial time series forecasting using rough sets with time-weighted rule voting. Expert Systems with Applications, 66, 219-233.
19. Qiu, M., Song, Y., & Akagi, F. (2016). Application of artificial neural network for the prediction of stock market returns: The case of the Japanese stock market. Chaos, Solitons & Fractals, 85, 1-7.
20. Shen, L., & Loh, H. T. (2004). Applying rough sets to market timing decisions. Decision support systems, 37(4), 583-597.
21. Shen, Q., & Jensen, R. (2007). Rough sets, their extensions and applications. International Journal of Automation and Computing, 4(3), 217-228.
22. Shyng, J. Y., Shieh, H. M., & Tzeng, G. H. (2011). Compactness rate as a rule selection index based on Rough Set Theory to improve data analysis for personal investment portfolios. Applied Soft Computing, 11(4), 3671-3679.
23. Skowron, A., Stepaniuk, J., & Swiniarski, R. (2012). Modeling rough granular computing based on approximation spaces. Information Sciences, 184(1), 20-43.
24. Suraj, Z. (2004). An introduction to rough set theory and its applications. ICENCO, Cairo, Egypt, 3, 80.
25. Tay, F. E., & Shen, L. (2002). Economic and financial prediction using rough sets model. European Journal of Operational Research, 141(3), 641-659.
26. Upadhyaya, S., Arora, A., & Jain, R. (2006). Rough Set Theory: Approach for Similarity Measure in Cluster Analysis. In DMIN (pp. 353-356).

تقدير عوائد الاسهم باستعمال نظرية المجموعة التقريبية: دراسة استكشافية مع دليل من سوق العراق للأوراق المالية

م.م. محمد حمدان عدنان
الجامعة التقنية الشمالية ، الموصل ، العراق
mohammed.hamdan@ntu.edu.iq

ا.د. مصطفى منير اسماعيل
جامعة بغداد, كلية الإدارة والاقتصاد ، بغداد ، العراق
Mustafa.m@coadec.uobaghdad.edu.iq

Received:28/4/2021 Accepted: 30/5/2021 Published: June/ 2021

هذا العمل مرخص تحت اتفاقية المشاع الابداعي نَسب المُصنَّف - غير تجاري - الترخيص العمومي الدولي 4.0
Attribution-NonCommercial 4.0 International (CC BY-NC 4.0)

المستخلص:
يهدف البحث إلى تقدير عوائد الأسهم وفق منهج نظرية المجموعات التقريبية واختبار فعالية اداءها ودقتها في التنبؤ بعوائد الأسهم وإمكانياتها في مجال الأسواق المالية وترشيد قرارات المستثمرين. وبلغت عينة البحث (10) شركات متداولة في سوق العراق للأوراق المالية, واظهرت النتائج قدرة كبيرة لنظرية المجموعات التقريبية في اختزال البيانات مما يساهم في ترشيد قرارات الاستثمار , وأبرز الاستنتاجات هي قدرة نظرية المجموعة التقريبية في التعامل مع البيانات المالية وتطبيقها لأغراض التنبؤ بعوائد الأسهم, ويرفد البحث المهتمين بالاستثمار في مجال الاسواق المالية والمستثمرين بأدوات تحليل مالي معتبرة تفوق الاساليب الاحصائية التقليدية وتكمن اصالة البحث في تنويع ادوات التحليل المالي والاحصائي واساليب التنبؤ بعوائد الاسهم.

المصطلحات الرئيسة للبحث: عوائد الاسهم , نظرية المجموعات التقريبية, السمات الشرطية, قاعدة القرار, التقريب للأقل, التقريب للأعلى.

*بحث مستل من أطروحة دكتوراه