Research Article

Providing an intelligent recommender system to increase the profitability of investors in the Tehran Stock Exchange

Edris Nanvai Savojbolaghy[a], Seyyed Mohammad Zargar[b]*, Seyyed Mohammad Ali Khatami Firouzabadi[c]

[a] Master of Information Technology Management, Islamic Azad University Electronic Branch, Tehran, Iran.
[b] Assistant Professor, Management of system, Department of management, Semnan Branch, Islamic Azad University, Semnan, Iran.
[c] Professor, Management and Accounting Department, Allameh Tabaraba'i University, Tehran, Iran.

Article History: Received: 5 April 2021; Accepted: 14 May 2021; Published online: 22 June 2021

Abstract: Recommender systems are the most important tools of data mining, which have been today created to predict the profitability in the stock exchange and providing relevant information. These systems can collect the data related to the stock exchange and the status per share from different references explicitly or implicitly and use them to predict and increase the earnings of investors. The investors used to apply return on investment (ROI) ratios to make financial decisions. Although the ratios gained pervasive function in practice, they never considered the time value of money (TVM) and investment risk. Later, the concept of TVM entered into the literature of the financial and investment economy and left significant effects on financial decisions. Accordingly, this study has proposed a recommender system to enhance the profitability of investors in the Tehran Stock Exchange. To do so, the combination of the NSGAII algorithm and clustering technique based on the K-Means algorithm was used. NSGAII algorithm can be used in this solution to derive the main specifications from the dataset to improve the efficiency and accuracy of clustering. In this case, the specifications with no impact on the prediction of profitability are eliminated, and the clustering is done on the main specifications without outliers. The process of elimination of outliers would be completed without losing the basic data of the dataset. Therefore, a purified set of rules is obtained by narrowing down the data, which can facilitate decision-making. In the end, the proposed solution was implemented in MATLAB based on Crisp Methodology. Using the solution, the risky, low-risk, and moderate-risk portfolios can be extracted. The purchase volume from the symbols of the five companies listed in the stock exchange differs based on the risk level and expected return.

Keywords: recommender systems, return on investment (ROI), Stock Exchange, NSGAII, K-Means

1. Introduction

Nowadays, due to the increasing growth of the Internet and a large volume of information, we require systems that can recommend the most appropriate services and products to user. One of the most important topics in cyberspace technology is the performance of recommender systems. The recommender system provides appropriate recommendations for the user by intelligently analyzing of information. Utility is a very important issue in a recommender system. In recommended systems, utility of an item is usually determined by pricing and it expresses how a particular user likes an item. In general, utility can be an optional function including the profit function. The utility can be defined as the rate defined by the user or by the program under the profit-based utility function. On the other hand, risk and return are basically two factors that have always been considered in the area of investment. Over time, risk versus return has become more comprehensive in nature and methods have been proposed to quantify measurement or investment risk management. Extensive recommender systems have been used to generate recommendations using different techniques in various application areas such as stock exchanges and this has created many challenges. The most important weakness of most techniques is the need to label features, resulting in its exclusive use in stock market intelligence. However, in the case of stock, description of new features is due to the lack of correspondence with the user's previous items. Thus, the user is unaware of the addition of new items to the system and limits the user's system recommendations to observing similar and duplicate items.

Recent studies have shown that in comparison between the recommendations of the stock exchange and the recommendations of the recommender systems in terms of quality and utility, the recommendations of the stock exchange are preferred, even if the recommendations provided by the recommender system include new and diverse factors. For this reason, integrating content computational models into recommender systems has become a research topic. The existing differences between growth and value stocks and the factors affecting them make investors invest in stocks by acquiring new financial knowledge and paying more attention to market conditions and periods in stocks (Perm Sankara, 2015). The issue of selecting portfolio is one of the difficult and complex issues for companies, organizations and even the general people involved with it, since successful predicting of portfolio promises profit and return on investment. The complexity of selecting the right portfolio is because profit of a stock depends on several factors, including political events, economic conditions, traders' expectations, and other environmental and social factors. These issues have made stock prices a dynamic, nonlinear, non-standard and irregular issue. Hui showed that in order to reduce investment risk and increase portfolio returns, the behavioral characteristics of stocks in previous periods should be examined to predict its future behavior. In addition, as stated before, one of the ways to reduce investment risk is to take advantage of the diversity in investing in creating a portfolio. Thus, in order to have the maximum return and minimum possible risk, we can select the stocks to form a portfolio among the stocks of various companies (Shackelford, 2017). Financial managers prefer to have a mechanism that can help them in their decision-making. For this reason, predicting methods has received much attention. Thus, capital market experts have examined the market for many years and identified different models for predicting. For this purpose, they have used a combination of model recognition and experience of observing cause and effect relationships (Kristoffersen et al., 2019). In addition, one of the most important issues in evaluating recommendation systems is related to the cold start issue. The cold start problem occurs when there is no appropriate evaluation for a new user who has recently entered the system, so no recommendations can be done for this user. The use of intelligent algorithms such as clustering to cover problem solving and reduce the error of recommender systems in the stock market can be useful, so an intelligent recommender system is provided in this study to increase the profitability of investors in the Tehran Stock Exchange using mixed clustering algorithm to cover several purposes in this system. Nasim et al. (2017) proposed an ontology-based recommender system for Commercial Off-The-Shelf (COTS) components. With development of COTS, it improves the identification of COTS components. It integrates these technologies into a comprehensive information retrieval technology and knowledge about COTS components and users to provide the most appropriate COTS components that meet the needs of users.

Ontology-based recommender system for COTS components is known as ONTOCOTS, which describes COTS components and makes their heterogeneous descriptions available on the web, and (2) a user model that represents user settings and interest domains. The proposed recommender system is divided into two main processes. The first process is responsible for extracting information about COTS components from the repositories and displaying them as ONTOCOTS items. The second is a descriptive process in which the user query is expanded using ontology WordNet and itis used to generate a formal query. The list of results is ranked based on the degree of satisfaction with the user's needs and preferences. Experiments suggest that the recommender system performed better in placing the relevant COTS components on the list. Jian et al. (2017) proposed a recommender system for solving intelligent stock exchange problems, using evolutionary evaluation methods. Instead of interacting with a small population, we form a very large initial population that is then divided into two categories. We used this approach in several evolutionary models without limitations and compared its performance with standard evolutionary algorithms performance. Using several evaluations (below 100), we obtained comparable results to advanced evolutionary algorithms. Before performing intelligent-based optimization tasks using traditional EAs, it is recommended to try other techniques, such as those discussed here, to achieve similar results with less evaluation. Don et al. (2017) introduced a stock exchange recommender system to help select project managers. In this method, different classification combinations and feature selection algorithms were first evaluated and the best one was selected to build a recommendation model with stock exchange data. Then, the recommended model was used to predict a new stock exchange project. In this work, the number of defects, the severity of defects and changes in the stock exchange were also considered. Experiments on data sets of 37 different development teams from different countries show that the accuracy of the recommended method is 82.5%. Each of the methods has several shortcomings that reduce the efficiency of current systems. For example, if a new item is added to the system, there is no way to recommend it to a user, unless another user rates it or introduces it as similar to other items. However, due to the ability to provide better recommendation, recent research works focus primarily on this technique and seeks ways to improve it. In the present study, we will try to cover more comprehensively than solving the problem space by presenting the clustering approach. Many studies have been conducted in different fields related to the subject of our research, including providing a recommender system and examining these systems, but the recommender systems have not been investigated so far using mixed clustering algorithms. In the present study, an intelligent recommender system was presented to increase the profitability of investors in the Tehran Stock Exchange. The present study also aims at answering these questions:

Which parameter in the objective function leads to an increase in the accuracy of the recommender system?
- What are the intelligence criteria of stock exchange for the clustering algorithm to solve the cold start problem?
- How the information of securities and stock exchange can be used in recommender systems to improve the accuracy and reduce the error of the recommender system?

Proposed solution

In the proposed method, before applying the data to the proposed model, the data preprocessing process is performed to normalize them. This section includes the following steps:
- Data homogenization (training and test data): In this section, a type of homogenization is performed on the data level by placing the letter characteristics of the data set with numerical values.
- Data normalization (training and test data): In this section, the imbalance between the data is eliminated. Since some features have large numerical values that can dominate other features, this section eliminates this group of abnormalities.

In the next step, the operation related to NSGAII algorithm is performed. In general, important features of data are extracted by using the NSGAII algorithm to increase the overall performance of the solution and efficiency of the clustering sector, since feature reduction can ultimately increase the efficiency of the clustering algorithm. In methods that lack this capability, the algorithm has to use features that have nothing to do with stock profitability. This type of learning will actually be learning with outliers and will have a negative impact on the recommender system. In the proposed solution, the clustering algorithm (K-Means) uses only the important features provided by the NSGAII algorithm instead of using all the features, resulting in an increase in the accuracy of solution. Moreover, processing, applying, and utilizing of features that do not play an important role in clustering reduce the clustering process and solution speed. In this method, redundant data is deleted without losing basic database data. Therefore, information reduction results in a concise and meaningful set of rules that facilitate decision-making and learning. In fact, the genetic algorithm maps the raw data space to the concept space by reducing the data space and selecting important features. In other words, in this case, the stock features are summarized into some of the most important features and increases the efficiency and accuracy of the stock status. The most important reasons for selecting NSGAII algorithm are the following cases:

- It has a faster solution than other methods and the problem of computational complexity of previous algorithms has been resolved in this method. Assuming that M is the number of target functions and N is the population size, the computational complexity of this method is O (MN^2), which is much less than the computational complexity of previous methods.

-It uses crowding distance to obtain a uniform set of solutions from other algorithms and estimate density of points around the solutions. In fact, crowding distance is a factor that is used to better select the solutions in terms of scattering on a set of solutions.

Accordingly, the details of each of the parts of the proposed solution are provided below.

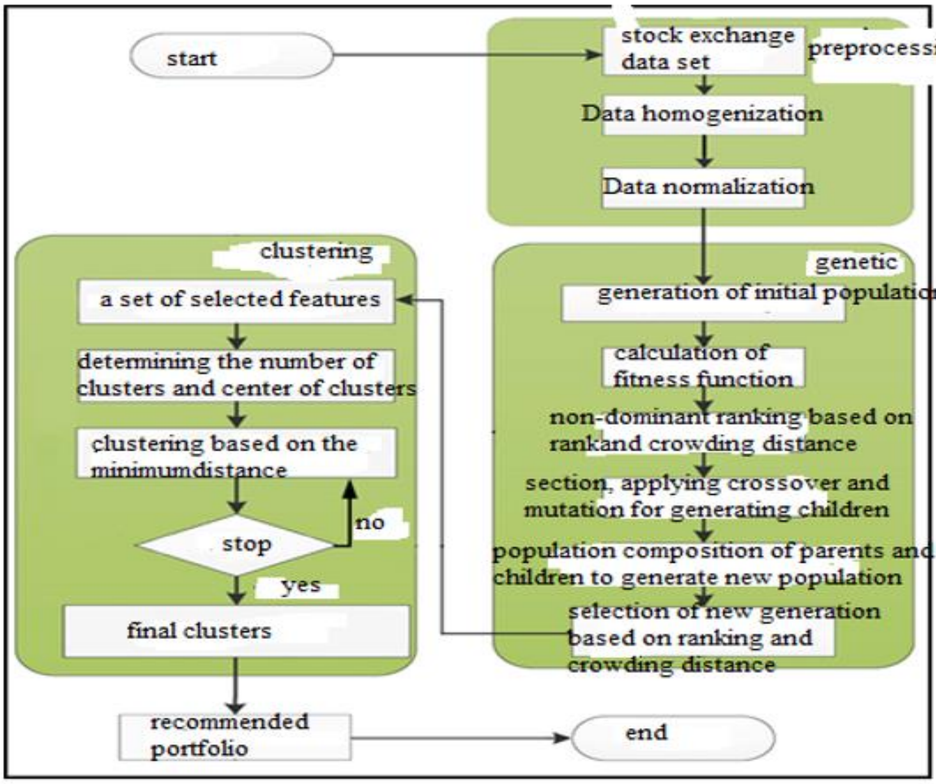

Figure 1- Proposed model

Non-dominated sorting genetic algorithm II (NSGA-II)

In the first step of NSGA-II algorithm, the initial population is generated. In the second step, a rank and a crowding distance are assigned to each member of the population. The method of calculating the rank is such that for each member i, two characteristics ni and Si are calculated. ni is the number of dominant members of i and Si are the set of recessive chromosomes i. The members with ni = 0 are the first Pareto frontier or fi. Now, for each member of fi, a member of the recessive set Si is considered and nj for the jth member of it is reduced by one unit. The member in which nj = 0 will belong to the H set. After completing H for all members of f_1, we can say that H is the second Pareto frontier. Then, f1 is set aside and H is used as the first Pareto frontier and the above process is repeated for the rest of the members .In this step, the crowding distance is calculated for each member. To determine the average, the crowding distance of the nearest member on both sides of the point is considered for all utility functions. The quantity of crowding distance indicates the size of the largest meta-rectangle that firstly takes a member and secondly does not include any other member. The following figure shows this concept for two utility functions.

In the third step, among the solutions of each generation, a number of them are selected using the tournament selection method. In the tournament selection method, two solutions are randomly selected from the population. Then, two solutions are compared and finally the superior solution is selected. The selection criteria in the NSGA-II algorithm are primarily the solution rank and secondly the swarm interval related to the solution. The solution that has a lower rank and a greater crowding distance is considered as desirable solution. In the fourth step, by repeating the tournament selection operator on the population of each generation, a set of members of that generation are selected to participate in the combination and mutation. The combination and mutation are performed on a part of the selected members and a population of members is created. In the fifth step, this population is merged with the initial population and the members of the new population are first sorted based on rank and in an ascending order. Then, the members of the population, who have the same rank, are sorted in a descending order in terms of distance and crowding. Then, the members of population are sorted first in terms of rank and then in terms of crowding distance. Finally, in the final step, the highest-ranking members of the list are selected as the next generation population. This cycle is repeated until the termination criteria are met.

Selection operator

To apply NSGAII operators to solutions (Cis), Cis must first be selected as the parent to create child solutions based on these solutions. In this solution, the tournament selection method is used to select the parent solutions. This method is one of the most widely used methods in genetic algorithm due to its efficiency and simplicity of implementation. In this method, n samples from a large set are randomly selected and the selected samples are compared with each other. The sample with the highest level of fitness wins and participates in the next generation. The number of samples entered in each competition depends on the size of tournament, and its value is usually assumed to be 2, which is also called as a binary tournament. Since diversity can reduce the rate of convergence, this method gives all the samples a chance to be selected.

Crossover operator

In this research, order-based crossover operator was used. A cut-off point is selected randomly. The first segment of the chromosome is transferred from the parent to the corresponding child without change and the other genes are transferred from left to right as they are in the other parent (non-corresponding parent) (second child is also obtained in the same way).

Mutation operator

In the present study, two mutation operators, including swap and inversion operators, were used. One of them is randomly selected and is applied on the desired solution. The swap operator takes the current solution and its two elements and creates the new solution by changing the location of these two elements. The inversion operator takes the current solution and the two elements of it and creates the new solution by changing the location of the two elements and sorting the middle elements in the opposite direction to the original solution elements.

Fitness function

The fitness function is a key component of the NSGA algorithm, used to evaluate usefulness or non-usefulness of one member to survive. This dissertation presents a new fitness function for the NSGA algorithm to reduce error rate and increase the rate of true positive (TP) in selecting useful features. Accordingly, in the fitness function, two new parameters called error rate (Error) and true positive rate (TPR) are used as follows:

$$\text{Fitness (S)} = W_a \text{TPR} + W_b \text{ Error}$$

The true positive rate parameter is equal to the ratio of the number of correctly selected samples to the total number of samples. This parameter is calculated in this way.

$$TPR = \frac{TP}{TP + FN}$$

The error rate is equal to the number of samples that have not been selected correctly. This parameter is calculated in this way:

$$Error = \frac{FP + FN}{TP + FN + FP + TN}$$

Clustering section

The K-Means algorithm is one of the simplest and fastest clustering methods. This algorithm has a parameter called k that specifies the number of clusters that must be obtained.

In the proposed solution, the K-Means clustering algorithm is implemented as follows:

Defining the number of k clusters: For example, if k = 2, we assume normal and abnormal traffic in training two different clusters.

- Initialization of k cluster centers of gravity, this is done by randomly selecting k data items from the data set.
- Calculating the distance of each item to the centers of gravity of all clusters using the Euclidean metric distance, which is used to find similarities between items in the data set.
- Assigning each item with the nearest center of gravity of the cluster: in this method, all items will be assigned to different clusters, so that each cluster will have items with similar features
- After assigning all items to different clusters, the mean of the modified clusters is recalculated. The recently calculated mean is assigned as the new center of gravity.
- Repeating step 3 until the center of gravity of the cluster does not change.
-Normal and abnormal labeling of clusters depends on the number of data items in each cluster.

However, it can be stated that there is a fundamental problem in this method of clustering. This problem is determining the initial part and the appropriate number of K clusters.

The overall process of solution

To provide a clear and accurate view of the study problem and the way of solving it, the proposed method was used for solving the problem of stock prediction in accordance with Crisp methodology. This methodology is in fact the standard industry process for Cross Industry Standard Process for Data Mining (CRISP-DM) and consists of six steps in a circular process.

The steps of this model as the method of implementing this research are as follows:

1- Problem definition

2- Data analysis: In this step, data of stock price of five companies are from more than 350 companies listed in the stock exchange as the sample of study. The required data (daily stock price of the selected companies) is uploaded from the website of the Tehran Stock Exchange Technology Management Company at TSETMC.com.

3- Data preparation: In this step, the data collected in the previous stage will be descriptively examined. The nature of data, the granularity of data and the quality of the data are considered in this stage. Data reduction will also be done in this step.

4- Modeling: In this step, data is divided into training data set and test data set. Since the volume of examined data is not very large, it is not necessary to allocate a part of data to validation data.

5-Evaluation: In this research, the strategy of selecting the portfolio will be evaluated based on risk and return parameters. Using different levels of risk and return and their ratio, the volume of purchase is determined in the portfolio.

6-Deployment: In this step of study, the proposed solution, which is based on the non-dominant genetic algorithm and K-Means clustering, is deployed in the MATLAB program environment

Data collection

Five companies from five heterogeneous industries were selected for this research as described in Table 1

Table 1- Symbol of selected companies

Row	Company	Trading symbol	Group
1	Minoo Shargh Food Industries	Ghaminoo	Food and drink
2	Khark Petrochemical	Shakharak	chemical products
3	Omid Taban Hoor Energy Management	Vehoor	Supply of electricity, gas, steam and hot water
4	Pars Khodro	Khpars	Automobile and Manufacturing automobile parts

5	Naft	Vanaft	Oil products, coke and fuel

The data resulting from daily transactions in the Tehran Stock Exchange are stored on the information servers of that organization and are provided in the form of web service or a downloadable file for storage to the interested people. Web service or more specifically the xml file in the information technology of the stock exchange and the daily data file are provided to legal entities only according to the rules and requirements of the stock exchange. The file can be downloaded and saved in csv and xls formats and can include the information fields of symbol title, first price, the highest price, the lowest price, final price, transaction volume, transaction value, and the last price. To download and store the daily transaction price data of the five selected stocks, the information technology site of the stock exchange was used and the desired data were stored in the form of a csv file and then converted to mat format for use in the MATLAB program.

2. Results

Daily stock price data are dynamic data type. The daily price Diagram of the stock of different symbols is shown in Figure 2.

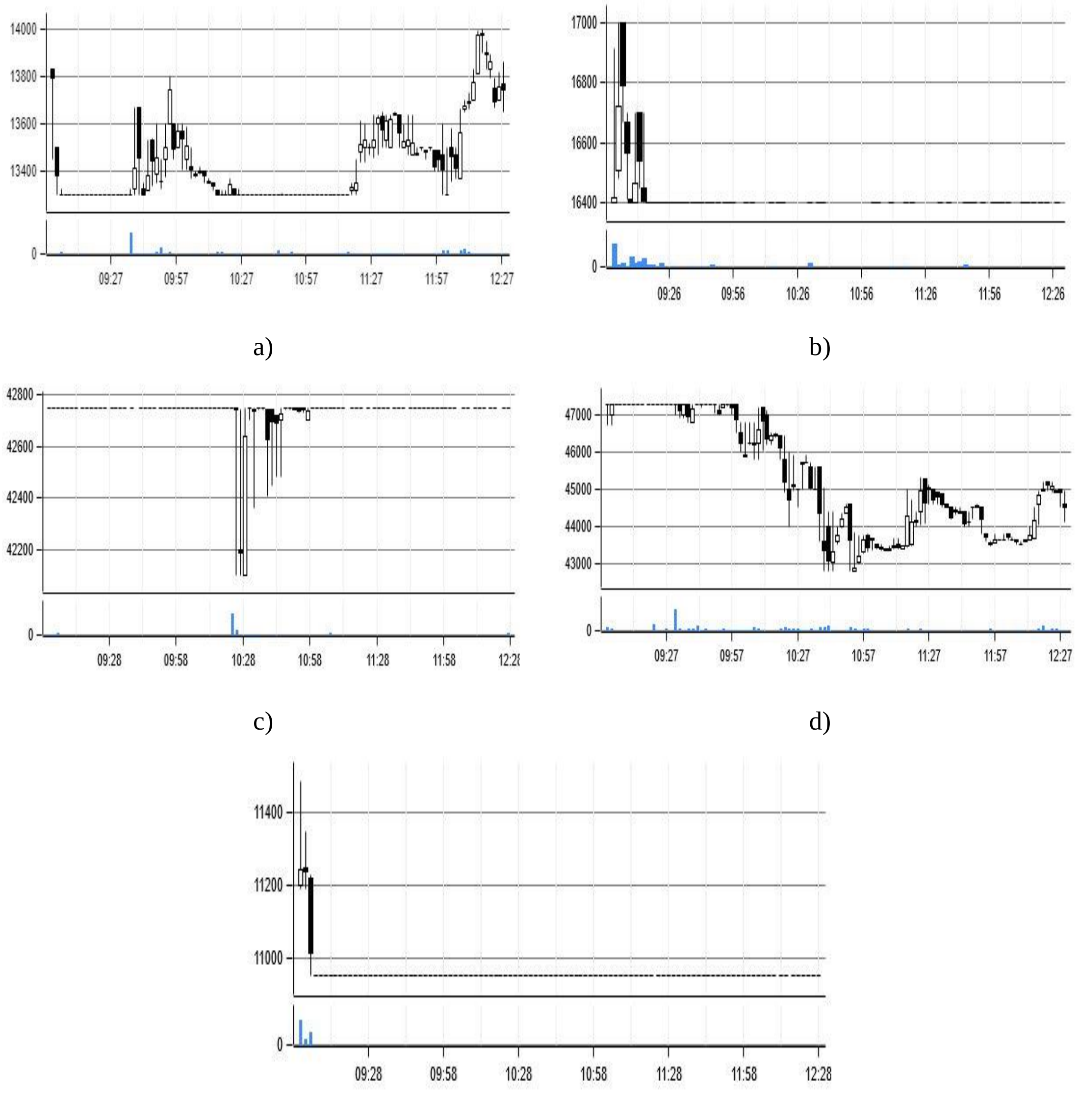

Figure 2 - aA) Daily stock price Diagram of Naft symbol, B) Daily stock price Diagram of Ghaminoo symbol, c) Daily stock price Diagram of Shakhark symbol, D) Daily stock price Diagram of Vehoor symbol, E) Daily stock price Diagram of Khpars symbol

Data quality

In the data analysis and preparation step, the most important task is to check the quality of data in terms of accuracy, completeness, non-redundancy, non-missing and non-disturbing. Stock price data are accurate data, because they inherently show the exact size and the judgment of the two observers on them is exactly same. In addition, since stock price data are sufficient for predicting, they are complete data. In examining the redundancy of stock price data, since there is no more than one variable in this study, redundancy caused by strong correlation of variables is rejected. Regarding the redundancy caused by unrelated data, since this solution uses the non-dominant genetic algorithm (NSGAII), concise and appropriate data are always extracted. In other words, by deleting irrelevant data, the size of data can be reduced and thus the productivity level of the solution can be increased. Figure 3 shows the price Diagrams of these five companies for the last 20 workdays.

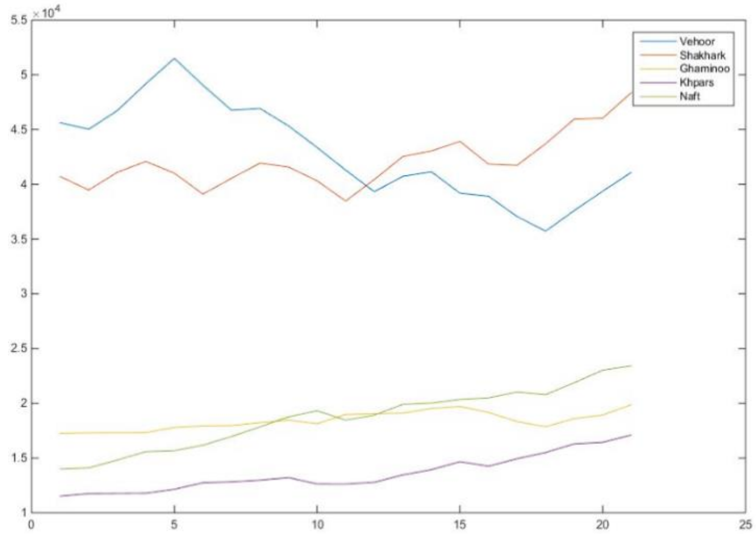

Figure 3 - Diagram of five selected companies based on their prices

As shown, stock price is not a good indicator for scaling and analysis, since the price of a stock may be several times higher than other prices. Accordingly, there is a need for an index that can examine all these prices together and analyze the rate of increase and reduction in a given period. For this purpose, the return index is used (Sias et al., 2020). This index actually shows how profitable a company is in relation to the total assets of that company. Accordingly, this index provides a general view of efficient management in relation to the use of assets to generate profits. It is calculated by dividing the annual profit by the total assets of each company.

Figure 4 shows a diagram of the five companies after scaling using the return index. With the help of this index, the scaling related to prices has been removed and the data of each company have been placed in the same scale along with other companies based on their return.

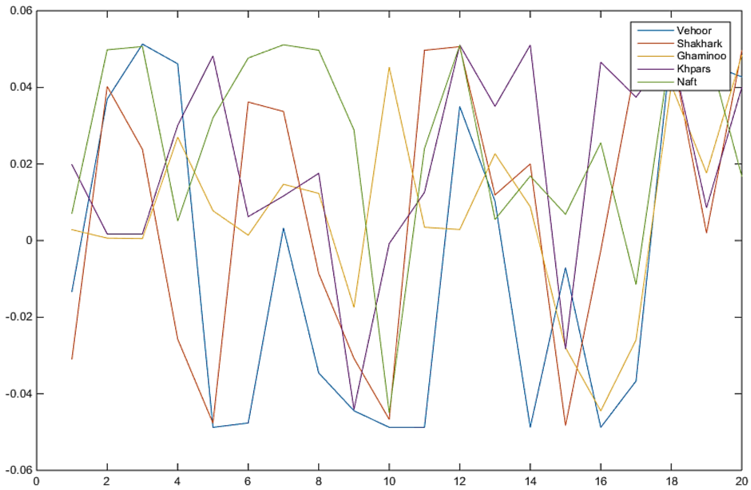

Figure 4) Diagram of five selected companies based on return index

To deploy this model and enter the data to NSGAII algorithm in the MATLAB program, a matrix is created for the desired number of data. In other words, the stock price data are transferred from the input matrix environment resulting from the time series to another matrix environment with dimensions n * m. The results of deployment of the solution are shown below.

Deployment results

Once the proposed solution was deployed on the data set, the results of its output are shown as a Pareto diagram in Figure 5. This diagram is based on the rate of return and risk per stock.

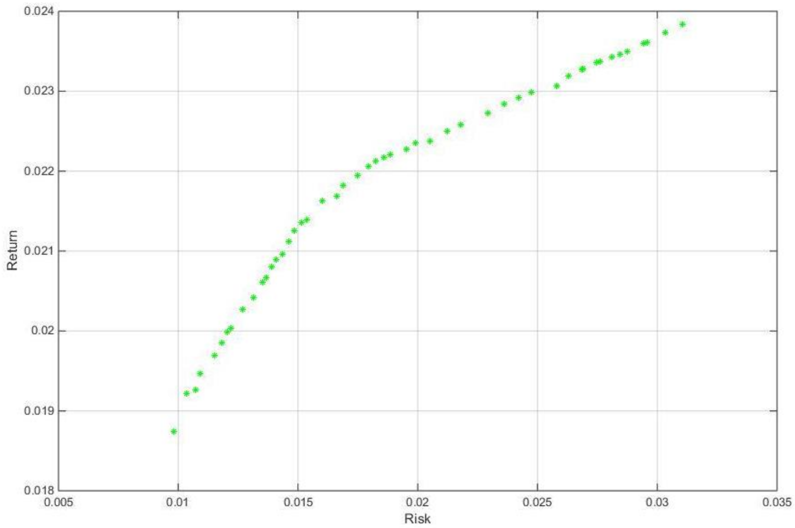

Figure 5) Output derived from deployment of the solution

As shown in Figure 5, the proposed solution has reversed a Pareto set. In this Pareto frontier, which represents the set of optimal solutions, the portfolio can be selected according to the type of strategy. In fact, the green diagram is related to the efficiency level. Each investment scenario either will be efficient on the curve in the best state or will be placed at the bottom of it.

Accordingly, were classified the results derived from output into three categories, including low-risk, high-risk, and medium-risk. The results are shown in Table 2. In these tables, the weight of each stock is shown according to the risk and return parameters. It means that if we want to have the maximum return, we must also accept the maximum risk, and vice versa. In fact, selecting the results derived from output of the solution allows the user to select the best portfolios based on aim of his investment.

Table 2) Portfolio of offered stocks with low, medium and high risk

Low risk	Return	Percentage of Vehoor symbol	Percentage of Shakharak symbol	Percentage of Ghaminoo symbol	Percentage of Khpars symbol	Percentage of Naft symbol
0.0109	0.0195	0	0	0.2558	0.2506	0.4935
0.0104	0.0193	0	0	0.2757	0.2186	0.5057
0.0101	0.0191	0	0	0.2952	0.1880	0.5167
0.0102	0.0191	0	0	0.2911	0.2911	0.5139
Medium risk	Return	Percentage of Vehoor symbol	Percentage of Shakharak symbol	Percentage of Ghaminoo symbol	Percentage of Khpars symbol	Percentage of Naft symbol
0.015	0.0213	0	0	0.0579	0.5608	0.3813
0.0135	0.0206	0	0	0.1309	0.4453	0.4238
0.0132	0.0205	0	0	0.1442	0.4251	0.4307
0.0124	0.0201	0	0	0.1862	0.3568	0.4570

High risk	Return	Percentage of Vehoor symbol	Percentage of Shakharak symbol	Percentage of Ghaminoo symbol	Percentage of Khpars symbol	Percentage of Naft symbol
0.0404	0.0251	0	0	0	0.1028	0.8972
0.0423	0.0254	0	0	0	0.0600	0.9400
0.0437	0.0256	0	0	0	0.0294	0.9706
0.0450	0.0257	0	0	0	0	100

As shown in Table 2, in the low risk state, only the purchase of stocks of Ghaminoo, Khpars and Vanaft symbols has been recommended. In practice, since these symbols are among the basic symbols in the stock market, purchasing stock of these symbols is less risky. In the low risk state, type of recommended portfolio and purchasing volume will also be different depending on the level of risk and return. Based on these four sets of solutions and the level of risk and return, no recommendation has been made of the symbols of Vehoor and Shakhark. In practice, as shown in diagrams of these symbols, they have had a descending trend, so it is possible that this trend will continue in the coming days. Accordingly, in this solution set, there is no recommendation to buy these two products. In the medium risk state, based on average risk, only basic stocks have been selected as a suitable option in the portfolio and symbols that have a low return on this level of risk have not been recommended. In the high-risk state, we consider returns at higher levels, so the risk also increases. In the latest solutions that have the highest risk, symbol of Vanaft has been recommended, since its return based on recent data sets will probably higher. As shown, with increasing the risk, the number of selected stocks will be lower, and the prediction gradually converges to a certain volume. In the highest risk state, it is recommended to buy only one stock with 100% volume, since it is possible to achieve maximum return compared to other solutions in this state, but the level of risk will be maximized.

3. Conclusion

The special importance of the capital market in economic development is undeniable through the effective management of capital and the optimal allocation of resources. Investment in the capital market requires decision-making, which in turn requires access to information about the future state of the stock market price. Thus, if the future trend of the stock market can be predicted with appropriate methods, the investor can maximize the return on his investment. In addition, the unknown factors influencing stock price changes are always a reason to predict companies' stock price. Nowadays, financial managers prefer to have a mechanism that can assist them in their decision-making, so prediction methods have received much attention. Hence, capital market experts have long studied the market and identifying different patterns for prediction. For this purpose, they have used a combination of pattern recognition and experience in observing cause-and-effect relationships. There are also many software programs that assist to this decision-making and are used as a predicting engine. However, in financial processes, some situations may arise that break the rules and make the prediction difficult by using these methods (Vismayaa et al., 2020).

The purchase of stocks of companies newly listed on the stock exchange is more risky compared to other listed companies, since investors do not know much about these companies. In other words, information asymmetry between managers and potential investors is the most important problem for the investors of these companies. In such a situation, companies newly listed on the stock exchange are forced to provide information that will help investors determine stock prices. Empirical evidence suggests that investors pay high attention to information such as earnings per stock prediction reported by managers when pricing stocks. These predictions reflect managers' expectations of the company's future performance. Therefore, the accuracy of these predictions is crucial for investors. Researchers in countries such as Australia, the United Kingdom, Hong Kong, Malaysia, New Zealand and Singapore have often studied the accuracy of earnings prediction reported by companies newly listed on the stock exchange. These countries have similar rules, regulations and institutions regarding securities. The situation in Iran is different from the situation of these countries. Therefore, it is necessary to identify the accuracy and role of predicted earnings in Iran. In Iran, state-owned economy is heavier than the mentioned countries, so that that Iran's economy can be called a state economy. However, in the recent years, based on the revision of Article 44 of the Constitution and development of a 20-year vision, the main goal of Iran’s economic officials is to reduce the government's economic activities and expand the private sector activities. To achieve this goal, it is necessary for Tehran Stock Exchange to be more active to pave the way for offering stocks of state-owned companies to public people.

However, in financial processes, often some situations arise that break the rules and make the prediction difficult by using these methods. In fact, stock price prediction in many cases can be very useful for investment organizations operating in various economic areas. In such a situation, w an intelligent system is required that can provide good results as the future price of stocks by receiving some data and training through data. Accordingly, the aim of present study was providing an efficient method in predicting stock price a using the NSGAII algorithm and K-Means-based clustering to increase the prediction accuracy. In this solution, the data processing process is performed to normalize them. In the next step, the operations related to the NSGAII algorithm are performed. In general, by using the NSGAII algorithm, important data features are extracted to increase the overall performance of the solution and to increase efficiency of the clustering, since feature reduction can ultimately increase the efficiency of the clustering algorithm .In the methods lack this ability, the algorithm has to use features that have nothing to do with stock profitability. This type of learning will actually be learning with outliers and will have a negative impact on the recommender system. Finally, through K-Means algorithm, stock clustering operations are performed to offer a series of portfolios to the user according to performance and risk parameters. The proposed solution was implemented in MATLAB program environment and in accordance with CRISP methodology, and high-risk, low-risk and medium-risk portfolios were extracted by using it. In each of these cases, based on the level of risk and the expected return, the volume of purchase from the symbols of the five companies listed in the stock exchange market volume will be different

References

1. Kristoffersen, E., Aremu, O. O., Blomsma, F., Mikalef, P., & Li, J. (2019, September). Exploring the Relationship between Data Science and Circular Economy: An Enhanced CRISP-DM Process Model. In Conference on e-Business, e-Services and e-Society (pp. 177-189). Springer, Cham.
2. Sias, R., Starks, L., & Turtle, H. J. (2020). Molecular Genetics, Risk Aversion, Return Perceptions, and Stock Market Participation (No. w27638). National Bureau of Economic Research.
3. Vismayaa, V., Pooja, K. R., Alekhya, A., Malavika, C. N., Nair, B. B., & Kumar, P. N. (2020). Classifier based stock trading recommender systems for stocks: An empirical evaluation. Computational Economics, 55(3), 901-923.
4. Chenkok, Sedigheh, Khatoon Habibi, Narjes, Norouzi, Alireza, 2016, Improving the accuracy of the proposed systems using the criterion of trust with similarity, First National Conference on Information Technology, Communications and Soft Computing, Isfahan, Islamic Azad University, Isfahan Branch.
5. Hassanvand, Reza, 2016, Weighing and identifying behavioral distortions of investors in the stock market, the 5th International Conference on Accounting and Management and the 2nd Conference on Entrepreneurship and Open Innovation, Tehran, Mehr Ishraq Conference.
6. Rahimian, Atieh, Yaghmaei, Farzin, 2016, improving emotion-based video recommending systems with the help of ontology, International Conference on Computer Engineering and Information Technology, Tehran, Permanent Secretariat of the Conference.
7. Rashidi, Atefeh, Shah Sanaei, Elham, Kaedi, Marjan, 2016, Design and implementation of a gift recommendation system based on user personality, First International Conference on New Research Achievements in Electrical and Computer Engineering, Tehran, International Confederation of Inventors the world.
8. Solgi, Ehsan; Hossein Ebrahimi Baghbanan and Pejman Mehran, 2015, Clustering of Growth and Value Stocks in Tehran Stock Exchange Using FCM Algorithm (Case Study of Companies Acting in Food, Beverage and Sugar Industry), 12th International Conference on Industrial Engineering, Tehran, Iranian Society of Industrial Engineering, Kharazmi University.
9. Kamyar and Pouya Shivoei, 2016, Modeling the Tehran Stock Exchange Market Index Forecast Using Particle Swarm Algorithm, 2nd International Conference on Accounting, Economics and Financial Management, Shahrekord, Payame Noor University, Shahrekord Branch.
10. Safavid, Seyed Rashed; Habibollah Hassanpour and Sedigheh Nazouri, 2016, Stock Price Forecast of Insurance Industry Investment Company in Tehran Stock Exchange Using ARIMA Model, 3rd International Conference on Economics, Management, Accounting with Value Creation Approach, Shiraz, Managers Training Institute Expert Narun.
11. Ghanbari Chah-e-Anjiri, Esmaeil, Honarvar, Alireza, 2016, improving the detection of profile injection attacks in recommender systems using remote point analysis in bulk data, First International Conference on New Research Achievements in Electrical and Computer Engineering, Tehran, International Confederation of World Inventors (IFIA), Comprehensive University of Applied Sciences.
12. Kajbaf, Mohsen, Rahmanian, Mohammad Sadegh, 2012, proposing an algorithm for clustering social network graphs based on central nodes, 11th National Conference on Intelligent Systems.

13. Golnataj, Zuleikha, Adelkhah, Maryam, 2016, Consulting Systems Based on Collaborative Refinement, Second National Conference on New Approaches in Computer and Electrical Engineering, Rudsar, Young and Elite Researchers Club.
14. Mohammadi, Pouria, Nasiri, Pegah, Shirmohammadi, Ali, 2016, Parallel Clustering Algorithm in Star Graph, International Conference on Computer Engineering and Information Technology, Tehran, Permanent Secretariat of the Conference.
15. Moradi, Gilda, Mahdavi, Mehregan, 2011, increasing the efficiency of e-commerce systems using recommending systems, First National Conference on Soft Computing and Information Technology, Mahshahr.
16. Mirshakari, Jamal and Khosrow Amirizadeh, 2017, Correlation-Based Clustering for Recommending Systems in E-Commerce Applications, The First National Conference on New Thoughts in Business Management, Tehran, University of Science and Culture.
17. Ismaili, Leila and Seyed Alireza Hashemi Golpayegani, 2016, a tourism advisory system in the context of social business: based on similarity, social communities, trust and reputation, 22nd Annual National Conference of Iran Computer Association, Tehran, Sharif University of Technology - Computer Association Iran.
18. Moradi Shoon, Gilda and Mehregan Mahdavi, 2010, Increasing the efficiency of e-commerce systems using referral systems, First National Conference on Soft Computing and Information Technology, Mahshahr, Islamic Azad University, Mahshahr Branch.
19. Yazdan Panahi, Behnam and Ahmad Mosalanejad, 2016, Application of Advisory Systems in E-Commerce, 5th International Conference on Computer, Electrical and Electronics Engineering, Kuala Lumpur - Malaysia, Pars Conference Institute, IUMW University of Kuala Lumpur.
20. Barth, M.E., Beaver, W.H. and W.R. Landsman."The Relevance of the Value Relevance Literature for Financial Accounting Standard Setting: Another View, Journal of Accounting and Economics, 2015.
21. C. PremSankara R.Vidyarajb K. Satheesh Kumar, Trust Based Stock Recommendation System – A Social Network Analysis Approach, ACM SIGSOFT Software Engineering Notes archive, Volume 38 Issue 4, July 2015, Pages 1-6.
22. Cappella, J. N., Yang, S., & Lee, S. (2015). Constructing Recommendation Systems for Effective Health Messages Using Content, Collaborative, and Hybrid Algorithms. The ANNALS of the American Academy of Political and Social Science, 659(1), 290-306.
23. Davoodi, E., Kianmehr, K., & Afsharchi, M. (2013). A semantic social network-based expert recommender system. Applied intelligence, 39(1), 1-13.
24. Gantner, Z., Rendle, S., Freudenthaler, C., & Schmidt-Thieme, L. (2011, October). MyMediaLite: A free recommender system library. In Proceedings of the fifth ACM conference on Recommender systems (pp. 305-308). ACM.
25. Ge, Y., Xiong, H., Tuzhilin, A., Xiao, K., Gruteser, M., & Pazzani, M. (2010, July). An energy-efficient mobile recommender system. In Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 899-908). ACM
26. Holthausen, R. and R. Watts,"The Relevance of the Value Relevance Literature," Journal of Accounting and Economics, 2011.
27. Jian fengChenVivekNairTimMenzies, Stock Constrained Recommendation in Tmall, Information and Software Technology, Volume 95, March 2016, Pages 281-294.
28. K. Lang, A novel stock trading prediction and recommendation system, Multimedia Tools and Applications, February 2018, Volume 77, Issue 4, pp 4203–4215.
29. Kazienko, P., Musial, K., & Kajdanowicz, T. (2011). Multidimensional social network in the social recommender system. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 41(4), 746-759.
30. Kothari, S.P,"Capital Market Research in Accounting," Journal of Accounting and Economics, 2016.
31. Mandar Ganesh Deshpande1, Ashwin Prabhu2, Nishi Tiku, Intelligent Recommender in Stock Market, International Workshop on Recommendation Systems for Software Engineering, 2007.Pages 68-72
32. NacimYanesa Sihem Ben Sassiab Henda Hajjami Ben Ghezala, Ontology-based recommender system for COTS components, Journal.