Specialized Information System for Support of the Process of Recruiting Securities

Taras Dubyniak[a], Oleksandra Manziy[b], Tomasz Gancarczyk[c], Andriy Senyk[b] and Yurii Futryk[b]

[a] *Ternopil Ivan Puluj National Technical University, Ruska str., 56, Ternopil, 46001, Ukraine*
[b] *Lviv Polytechnic National University, Stepan Bandera str., 12, Lviv, 79013, Ukraine*
[c] *University of Bielsko-Biala, Willowa St. 2, Bielsko-Biala, 43-300, Poland*

Abstract
The development technologies and functional capabilities of the information system created by the authors aimed at decision-making support of the formation of a set of securities, which enables potential investors with little experience to assess independently the effectiveness of the investment portfolio by simulating the dynamics growth of assets available on the financial market are described. The proposed information system uses visualization methods that present available tabular information in the structured form of schemes, graphs, and charts. The web-oriented solution provides an opportunity to analyze and forecast portfolios in real time based on the available types of shares of various companies.

Keywords 1
Information and communication technologies, mathematical methods, visualization, risk, data science, Python

1. Introduction

Due to the significant increase in interest in investment activities not only by specialized organizations and professional investors, but also by private individuals, there is a need for accessible instrument for monitoring the financial securities market. Specialized software products for financial risk management by means of carrying out the in-depth analysis, generating reports and investment scenarios simulation are available in the Internet.

It is known that most of information faced by the investor is presented in ordered tabular format, and according to the cognitive methodology, a person perceives visualized methods of presenting information. An accessible presentation of information about financial product makes it possible to assess by the consumer whether the chosen asset meets the needs and or the consumer is ready to accept the risks inherent in such a product. All this confirms the relevance of creating the information system to support decision-making for the formation of securities portfolio available for non-professional investor.

2. Analysis of available researches

At present, the investments in securities are one of the priority areas of any financial market, both within separate developed country and at the global level. The proposed MPT portfolio theory (Modern Portfolio Theory) [1] consists in the diversification by means of weakly correlated assets [2]. One of

CITI'2023: 1st International Workshop on Computer Information Technologies in Industry 4.0, June 14–16, 2023, Ternopil, Ukraine
EMAIL: d_taras@ukr.net (A. 1); oleksandra.s.manzii@lpnu.ua (A. 2); tgan@ath.bielsko.pl (A3); andrij.p.senyk@lpnu.ua (A. 4); yurii.futryk.mpmkm.2021@lpnu.ua (A. 5)
ORCID: 0000-0003-1529-6951 (A. 1); 0000-0002-6480-2307 (A. 2); 0000-0002-9709-0860 (A.3) 0000-0002-1614-512X (A. 4); 0000-0001-5271-9883 (A. 5)

the means of effective optimization of portfolio solution is the application of mathematical methods and information and communication technologies [3-5].

Analysis of information makes it possible to eliminate the need of making assumptions while taking important decisions and instead to develop investment strategies with greater confidence. A similar toolkit of financial market analysis and effective optimization of work with securities is used in the investment activity assessment [6, 7].

Such information and analytical systems include:

1. Riskalyze platform (https://www.riskalyze.com) provides tools for investment risk analysis, plan execution, creation and execution of investment portfolios as the service to financial advisors in the USA. Riskalyze platform is flexible, making it possible for the users to connect and integrate with any budget in real-time on any device. Due to easy dynamic interface and adaptive design the users are able to create and view their portfolios and mobile reports to them.

2. Risk management system Arbor Portfolio Manager (https://arborfs.com) is the portfolio, asset and fund management software that provides institutional investors with detailed trade and position management. The system includes in-depth portfolio analysis, as well as extensive reporting on clear activity on own funds.

3. The platform LogicManager (https://www.logicmanager.com) is the solution for SaaS (Software as a service) that supports multiple users making it possible to build it within five working days. This platform serves many industries which require the hidden risk management system. The platform LogicManager enables organizations to make decisions based on data comparisons and set goals. The programs include enterprise risk management, IT management, compliance management, external risk management and financial reporting compliance. The software main functionalities, which extend to different areas of decisions are: detection, assessment, mitigation, monitoring and reporting of risk.

4. The platform CammsRisk (https://cammsgroup.com) is software for comprehensive integrated approach to the management of risks, incidents, audits and portfolio filling. This solution can be used by personnel for the combination of all the requirements of risk management and their forecasting.

5. The platform Sharesight (https://www.sharesight.com) is awarded investment portfolio tracker used by thousands of self-employed investors and financial professionals. Sharesight uses 20-year historical data and is synchronized with brokers for for automatic tracking of trades, dividends and corporate actions. The app also makes it possible for the customers to get online access to all the data required for their tax reporting, including foreign investment calculations.

6. The platform Looker (https://looker.com) is browser-based and cloud-based business analytics platform designed for data investigation and analysis. This platform helps corporations collect and analyze data and then make considered decisions taking into account comprehensive risks.

7. The platform FundCount (https://fundcount.com) is investment and portfolio accounting software that tracks, analyzes and reports the value of composite investments in different portfolios. The software supports a wide range of asset types and portfolio evaluation methodologies. The environment enables you to create flexible reports according to user requirements, and carry out analytical research on available data.

8. The platform HiddenLevers platform (https://www.hiddenlevers.com) is risk technology analysis platform that provides safe investment management. HiddenLevers offers solutions for individual users and companies that are designed primarily for executives, financial advisors, asset managers and portfolio managers.

3. Description of the created information system

Owing to the significant increase of interest in the investment activities not only by specialized organizations and professional investors, but also by private individuals, there is the need for accessible instrument for monitoring the financial securities market [8-11]. In this paper, the information system for decision-making support for the formation of securities portfolio, accessible to non-professional investor is proposed. Based on the well-known rule that visually presented information is perceived by users better than organized tabular information [12-14], the developed

system provides the opportunity to carry out comparative analysis of the dynamics of changes in the value of shares of various companies in the form of graphs and charts.

The main page of the presented platform provides introductory information for beginner, which will help to understand how to weigh correctly the risk and make decisions during the formation of the investment portfolio. The main theses described on the initial page, will promt the user how exactly to reduce risks and manage the investment process.

The novelty of the presented information system for decision-making support for the securities portfolio formation is the possibility to carry out visual comparison of the assets available on the financial market in order to fill the investment portfolio with the corresponding equity participation. The service is aimed at Ukrainian audience, and the interface of the application program enables the customer, to determine the price of shares on the basis of historical data of the value of assets and to create a set of assets based on the forecasts offered in the system.

A. The process of information system creation

In order to select the working programming language of the information system, the review and analysis of the capabilities of modern information technologies and programming languages Python, R, C#, GoLang(https://towardsdatascience.com) is carried out and the decision to choose Python to perform the set task is made. It is due to the powerful libraries that Python [15, 16] provides predominant opportunities in the development of software application while working with investment risks. Additionally, the interactive JupyterNotebook web environment can also be used for faster data analysis and visualization. It is also decided to deploy the product on the basis of Dash web framework by Plotly company, which provides interactive web environment for writing web-oriented system, and also provides a powerful toolkit for visualization and construction of dynamic graphs. In turn, this tool provides flexibility and interactivity to the system during the use. During the creation of the information system, mathematical and graphic Python libraries, such as: Pandas, Matplotlib, Seaborn, Numpy, Datetime, Plotly, Scipy, Statsmodels, Sklearn, Pathlib are used. These libraries provide a wide range of tools for data analysis, visualization and machine learning. In particular, Pandas provides convenient interface for data processing that makes them easy to read, process, and store. Numpy makes it possible to work with multidimensional arrays of data, Matplotlib and Seaborn provide tools for visualizing data in the form of graphs and charts, which makes it easy to analyze data and find useful information. Scipy and Statsmodels provide tools for statistical analysis of data and performing various statistical tests. Sklearn provides machine learning tools that enables you to develop machine learning models, perform classification and prediction. And in turn, Pathlib simplifies the work with file system.

For convenient and fast data analysis and their visualization, JupyterNotebook included in Anaconda distribution is used. This interactive web environment makes it possible to work with code and data in a user-friendly format quickly. Input data are downloaded in .csv format for information processing in Jupyter environment. The processed data are automatically integrated into PyCharm environment in the form of python code. For visualization of the data characterized by the companies shares, the capabilities of the Pandas libraries, Matplotlib and the data visualization library on the Dash web framework from the Plotly company are used. Dash is a new, easy-to-use Python framework for building dynamic web applications. It is built on top of Flask, Plotly.js, ReactJs. This product makes it possible to create dashboards using pure code in Python programming language. Dash has open source and stores real-time statistics and price forecasts for selected assets on web pages. The application of these libraries helps make charts dynamic and enables you to store information about selected assets in user-friendly and readable format. In general, application of Jupyter Notebook, PyCharm, Pandas, Matplotlib, and the data visualization library on Dash web framework makes it possible to work efficiently with Big Data, carry out analysis and visualization of results, and even create dynamic web applications based on pure Python programming language code. Application of these tools in data science projects can increase work efficiency, simplify the development process, and save data researchers time and effort. It is also important to note that these tools have active community of users, which allows you to get support and help in solving problems during the projects development. In general, the use of such technologies makes it possible to expand the opportunities of data researches and provide more accurate results and forecasts in the field of finance and other industries.

In order to carry out data analysis, the function which downloads stock data from the information source of the financial information provider YahooFinance (https://finance.yahoo.com) by means of Pandas library in .csv format is created. It should be noted that this type of data is available from many sources, such as Yahoo Finance, Google Finance, Quandl. In the course of work the given APIs are compared, and it is concluded that the use of Yahoo Finance in combination with Python libraries is the best way. API Yahoo is the gold standard for API stock data, used by both individual and corporate users. This source quickly and conveniently uploads data in real time. To download financial data from Yahoo Finance, we can use the pandas_datareader library, which enables us to receive conveniently and quickly the data in real time. Moreover, we can use yahoo_finance library to retrieve financial data from this source. To access the data, we can use the function that downloads stock data from Yahoo Finance by means of Pandas in CSV format. This makes it possible to store and process conveniently the data from the previous session. The application of Jupyter Notebook and PyCharm enables us to work efficiently with Pandas and other Python libraries for data analysis and visualization. For example, due to Matplotlib we can display graphs and charts for more effective data understanding and drawing the conclusions.

Furthermore, application of data visualization library in Dash web framework makes it possible to create interactive and dynamic data visualizations that can be useful while. presenting the results of data analysis. Dash is an open Python library for creating web applications with high-quality graphical interfaces that supports interactivity, animation, and dynamic data change. It runs on Flask and React, and enables developers to create quickly web applications that can be accessed in any web browser. Dash provides high-quality graphical interfaces using Python programmer-friendly syntax. The library has many built-in components for creating graphs, tables, forms, interactive control elements that can be easily easily configured and modified to display data by request. The interactivity and dynamic change of the data makes it possible to create data visualizations that can be easily understood and enjoyable for the user. Additionally, Dash enables to add functionality to data visualizations, such as filtering and sorting, allowing users to gain additional information and make conclusions based on real-time data display.

The created product is a web-based system for risk analysis in theinvestment portfolio of shares or cryptocurrencies.

B. Application of the information system for data analysis

Data from the stock market such as some securities of large technical companies Apple (https://www.apple.com); Google (https://www.google.com); Microsoft (https://www.microsoft.com); Amazon (https://www.amazon.com); Tesla (https://www.tesla.com); Oracle (https://www.oracle.com) are selected for the development and testing of the information system.

The user has the opportunity to assess visually the shares growth dynamics for selected technical companies (Figure 1) as well as analyze the given charts taking into account the moving average overlay.

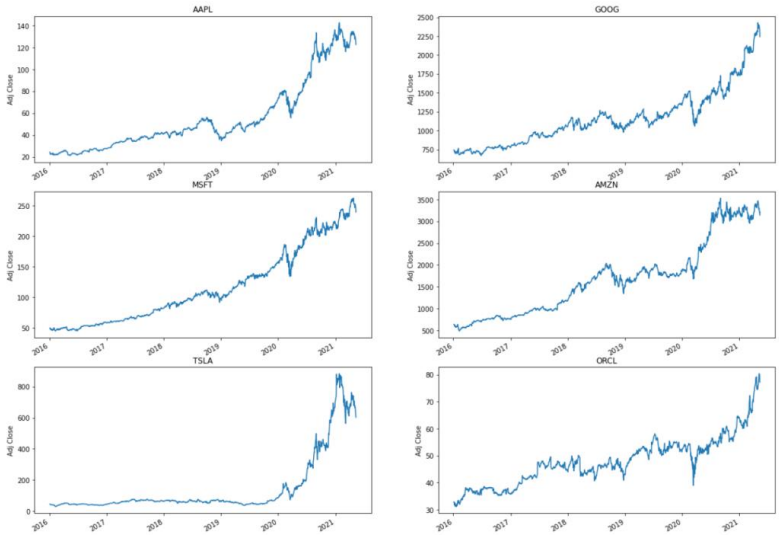

Figure 1: Growth dynamics for 6 companies in recent years

Besides the basic analysis, the system provides the opportunity to visualize the risk of lack of profit of each asset based on the analysis of the daily trend of share price changes. Graphs of the percentage change of daily returns are constructed using pct_change() on AdjClose(adjusted stock closing price) data column. While compiling the optimal tools portfolio, the indicator of tools correlation with each other, as well as their volatility, are used [17]. On the basis of analysis of the lines shown on the graph (Figure 2) for each asset as a separate part of the total investment portfolio, it is possible to predict the movement of the share price based on expected profits and risks.

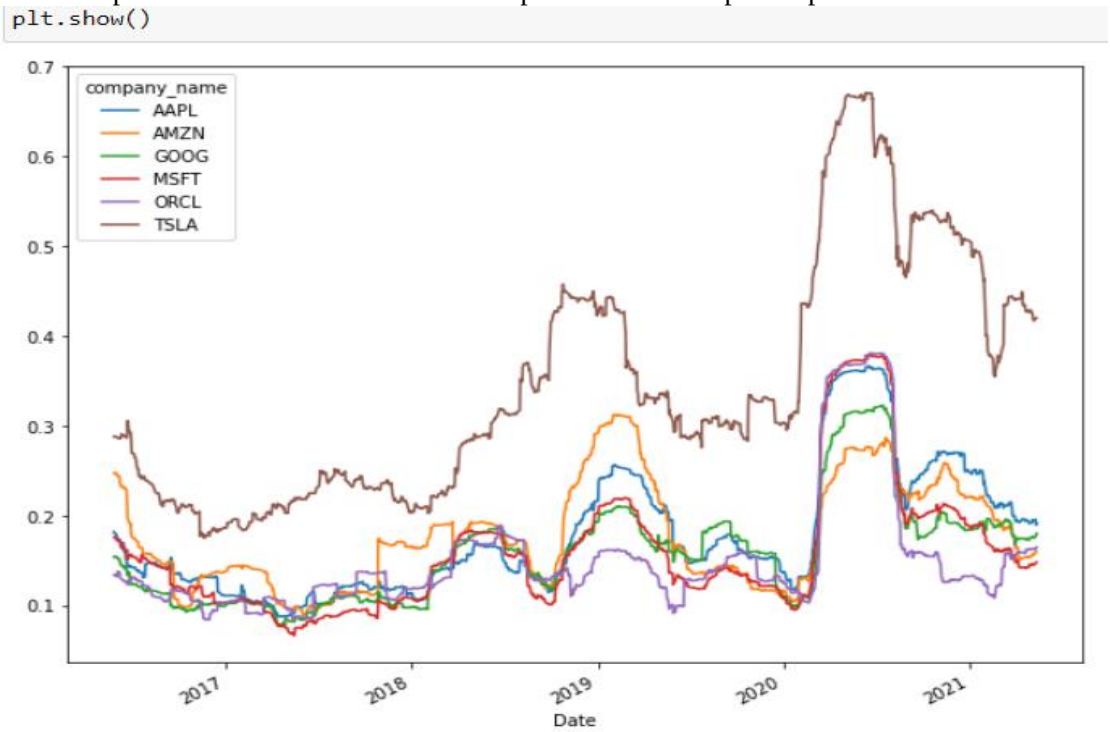

Figure 2: The graph of selected shares volatility

- Apple will bring a large income in comparison with others, but it is quite a risky stock.
- Google will bring less profit, but it is less risky.
- Microsoft will bring little profit, but the risk is sufficiently small.
- Amazon will bring less profit, but more risky than Microsoft.
- Tesla will bring the maximum return compared to other companies, but it is the most risky stock.
- Oracle will give the least profit, but it is the least risky investment.

4. Overview of the results of the information system operation

An important functionality of the system is the ability to forecast the price of financial instrument.

ARIMA stationary series forecasting methods are used, while two the most common stationarity testing methods are implemented that is visualization and Dickey-Fuller (ADF) test [18].

The evaluation metric used to search the set of parameters, the so-called grid, is AIC (Akaike Information Criterion) value. The resulting model clearly recorded the seasonality as well as the upward trend in the share price. We consider it to be a good result, because the average absolute percentage error gave the low percentage of error. This metric is a measure of accuracy of the methods performing forecasting in statistics. Due to it, we obtained the result of data forecast error (Figure 3).

Figure 3: ARIMA time series forecasting. Comparison of forecasts with actual prices

Since the system downloads data from the arbitrary web information source of open access, for better analysis of the investment portfolio, it is worth adding the comparison with more risky investments - cryptocurrencies, such as Bitcoin and Ethereum, which have shown a large increase since the end of 2020. Therefore, in order to determine the overall riskiness of investments, we will create the portfolio from a set of the above-mentioned stocks and cryptocurrencies.

Cryptocurrencies Bitcoin and Ethereum have the highest returns, but also the highest volatility, as expected, because the cryptocurrency market is quite volatile in comparison with the shares or securities market.

While calculating the efficiency of portfolio, Sharpe ratio is used, which is the ratio of the expected excess return of the portfolio to its volatility.

$$SR = \frac{E[R] - R_f}{\sigma} = \frac{APR - R_f}{\sigma}$$

The system provides the opportunity to create different sets of portfolios in order to select the best of them for further analysis.

The next stage of the investigation is analysis of the ratio of profit and risk. An effective representation of the result is their visualization (Figure 4), where the data points (portfolios) are colored based on the intensity of the corresponding color (the greater is the saturation of the tone, the denser are the predicted results are).

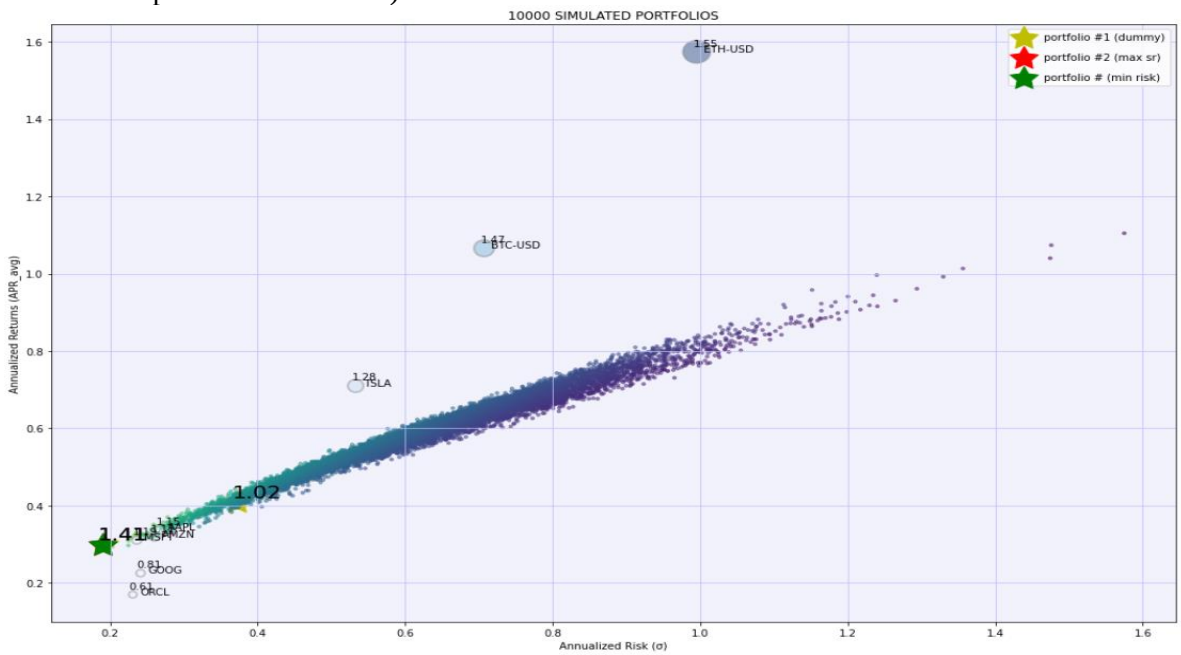

Figure 4: Application of Monte Carlo method for investment portfolio modeling

According to the results of technical and information analysis, we conclude that investing in the diversified portfolio is better option than investing individually in the risk-free asset or in individual instruments. Data analysis shows that selected cryptocurrencies increase the riskiness of our portfolio. Visualization of the optimal portfolio selection and calculation of its profitability and distribution of assets in the portfolio are shown in Figure 5.

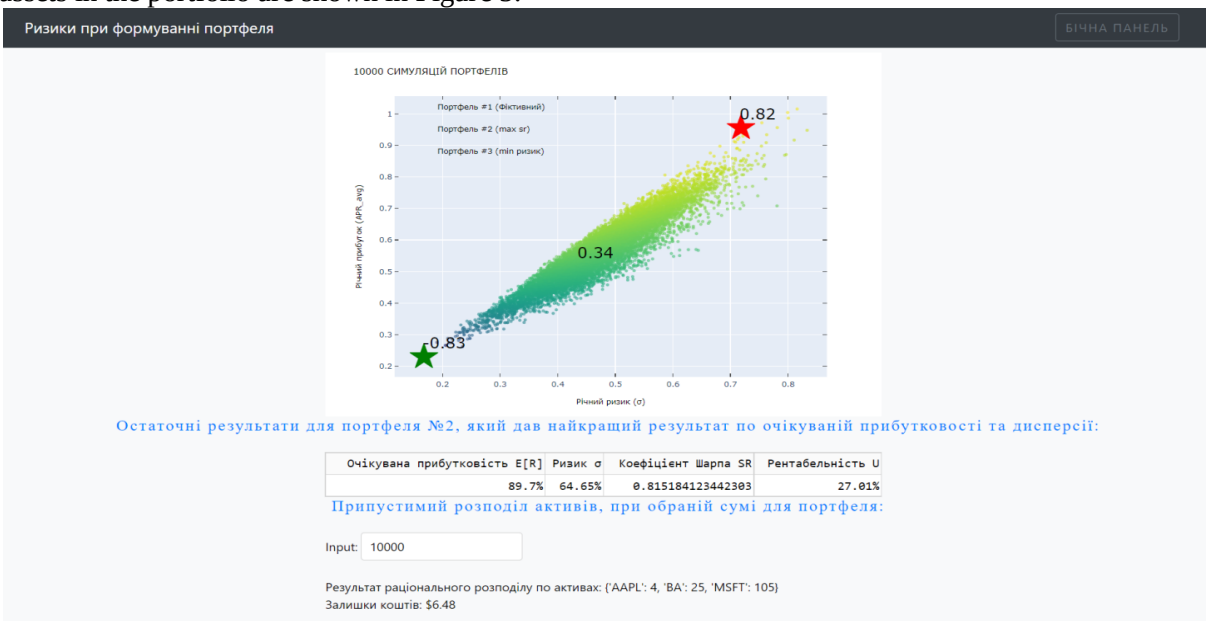

Figure 5: Updated expected returns and risk for the selected investment portfolio

As the result, the optimal filling of the investment portfolio is visualized in Figure 6 in the form of the diagram.

Figure 6: Diagram of the result of rational distribution of portfolio assets

The web-oriented solution provides the opportunity to analyze and forecast portfolios in real time by the available types of shares of various companies.

5. Conclusion

An overview of existing systems of the leading specialized data visualization and business intelligence software products used to analyze large volumes of data is offered. The global idea is designing computer algorithms, and computational informational high-level systems, which aim to minimize risk and maximize return based on the historical performance of the financial data. Furthermore, dynamic portfolio optimization and diversification are also considered as the target for further research that allows the designing of a flexible informational system to make profit maximization. This paper described the functionality and used technologies of the created information system. The visualization and forecasting methods used for analyzing investments in the financial

market in different time intervals, allow to support the dynamic diversification for different sets of financial assets. A collection of investment assets management refers to the process of investment decision making based on customized tactical investment strategies to match and maximize the return for each investing time horizon. The article has illustrated the usage of the proposed information system for analysis and forecasting while supporting dynamic diversification, of the investment process and to obtain the optimal set of financial assets in the selected time frame. The created information system is proposed to use as an advisory tool for individual non-professional or inexperienced investors with low financial stability. The results of the information system demonstrate that individuals who do not own significant finances can easily invest in a well-diversified set of financial assets, even if the risks are not fully diversified. When using a limited set of assets and a strict restriction on diversification, the results show that the application of such information systems is efficient and profitable. The following Python libraries were used in the process of complete creation for this information system: Pandas, Matplotlib, Seaborn, Numpy, Datetime, Plotly, Scipy, Statsmodels, Sklearn, Pathlib. In general, the use of such technologies as Jupyter Notebook, PyCharm, Pandas, and the Dash web framework for data visualization made it possible to efficiently work with large volumes of data to perform analysis and dynamic visualization of investment results.

6. References

[1] H. Markowitz, Portfolio selection, Journal of Finance 7(1) (1952) 77–91, https://doi.org/10.1111/j.1540-6261.1952.tb01525.x

[2] Kuzmin O., Alekseev I., Kolisnyk M. Problems of financial and credit regulation of innovative development of production and economic structures: monograph . Lviv Polytechnic National University Publishing House, 2007. - 152 p.

[3] J. Lu, D. RuanandG. Zhang (eds.), E-ServiceIntelligence: Methodologies, TechnologiesandApplications (Springer-Verlag, Berlin, Heidelberg, 2007), https://dl.acm.org/doi/10.5555/1203907.

[4] T. Stoilov, How to integrate complex optimal data processing in information services ininternet, in Proc. 20th Int. Conf. Computer Systems and Technologies, ACM DigitalLibrary, 2019, pp. 19–30, https://doi.org/10.1145/3345252.3345254.

[5] V. D. Ta, C. M. Liu and D. A. Tadesse, Portfolio optimization-based stock predictionusing long-short term memory network in quantitative trading, Applied Sciences 10(2020) 437, https://doi.org/10.3390/app10020437.

[6] Kalnyi, S. V. and Vysotskyi, V. A. (2019), "Management formation of investment portfolio enterprises in Ukraine", Efektyvna ekonomika, [Online], vol. 3, available at: http://www.economy.nayka.com.ua/?op=1&z=6953. https://doi:10.32702/2307-2105-2019.3.39

[7] Medynska, Tetyana V., Rushchyshyn, Nadiia M., and Nikonenko, Uliana M. (2020) "Tax Regulation of Investment Activity of Ukrainian Banks.", Business Inform 11:316–324. https://doi.org/10.32983/2222-4459-2020-11-316-324

[8] M. García-Galicia, A. A. Carsteanu and J. B. Clempner, Continuous-time mean varianceportfolio with transaction costs: A proximal approach involving time penalization,International Journal of General Systems 48(2) (2019) 91–111, https://doi: 10.1080/03081079.2018. 1522306.

[9] X. Huang and X. Wang, Portfolio investment with options based on uncertainty theory,International Journal of Information Technology & Decision Making 18 (2019) 929-952,https://doi.org/10.1142/S0219622019500159.

[10] E. Allaj, The Black–Litterman model and views from a reverse optimization procedure:An out-of-sample performance evaluation, Computational Management Science 17(2020) 465–492, https://doi.org/10.1007/s10287-020-00373-6.

[11] A. Palczewski and J. Palczewski, Black–Litterman model for continuous distributions,European Journal of Operational Research 273(2) (2019) 708–720, https://doi:10.1016/j.ejor.2018.08.013,
https://www.sciencedirect.comscience/article/pii/S0377221718306933.

[12] A. Rutkowska and M. Bartkowiak, Exertion approach to vague information in portfolioselection problem with many views, 2019 Conf. Int. Fuzzy Systems Association and theEuropean Society for Fuzzy Logic and Technology (EUSFLAT 2019) (Atlantis Press,Paris, France, 2019), pp. 142–149, https://www.atlantis-press.comproceedings/eus°at-19/125914792.

[13] F. Wen, L. Xu, G. Ouyang and G. Kou, Retail investor attention and stock price crashrisk: Evidence from China, Journal of International Review of Financial Analysis 65(2019) 1–15, https://doi.org/10.1016/j.irfa.2019.101376.

[14] G. Kou, Ö. Akdeniz, H. Dinçer and S. Yüksel, Fintech investments in European banks: Ahybrid IT2 fuzzy multidimensional decision-making approach, Journal of Financial Innovation 7(39) (2021) 1–28, https://doi.org/10.1186/s40854-021-00256-y.

[15] Wes McKinney. Python for Data Analysis / Wes McKinney, Julie Steele and Meghan Blanchette. – United States of America: O'Reilly, 2018. – 470 c.

[16] Jake VanderPlas. Python Data Science Handbook.Essential Tools for Working with Data / Jake VanderPlas. – United States of America: O'Reilly Media, Inc., 1005. Gravenstein Highway North, Sebastopol, CA 95472., 2017. – 548 c

[17] YaroshkoS., ManziyO. Financial mathematics. Part 1.Lviv, ZUKC Publishing House, 2021. - 209 p.

[18] Dickey, D. A.; Fuller, W. A. (1979).Distribution of the Estimators for Autoregressive Time Series with a Unit Root. Journal of the American Statistical Association 74 (366): 427–431. JSTOR2286348.https://doi:10.1080/01621459.1979.10482531