(MIJ) 2022, Vol. No. 8, Jan-Dec **e-ISSN: 2454-924X; p-ISSN: 2454-8103**

Developing a Financial Data Analysis in Conjunction with Risk Quantification Based on Python Programming Tools

Saniya Malik

Narayana e-Techno School, South City ll, Gurugram, India

ABSTRACT

Internet technology plays an increasingly significant role in collecting and analysing financial data as the level of information in the financial sector continues to rise. Internet technology enables accurate data collection and analysis on a large scale. The large-capacity financial database demands more efficiency and accuracy of data analysis tools. Python, a brand-new programming language, has the potential to become an essential tool for financial data collection and analysis by completing the sniffing and collection of Internet resources more effectively. Python technology faces several challenges and dangers due to scientific and technological advancements. The purpose of this article is to discuss how to analyse risks quantitatively using Python and the most important applications of Python in financial data analysis.

INTRODUCTION

Theoretical and practical circles have been paying more and more attention to quantitative analysis in the financial sector. Quantitative financial analysis combines computer programming and financial analysis theories. It also makes better use of modern computing technology to find trading opportunities and accurately price financial assets. The technology of quantitative analysis has also made significant progress and has become a hot topic of discussion. Currently, quantitative analysis encompasses all aspects of the financial industry, including quantitative investment, risk management, and pricing of basic and derivative financial assets.

People have begun to record bits and pieces of life in the form of data in recent years due to the rapid advancement of science and technology. These enormous amounts of data have practical value if processed and analysed. The financial industry uses more complex financial data than other industries, and some are random, making data analysis more difficult. As a result, one of the current scientific development directions is applying data mining technology to extract information concealed within massive amounts of data.

ADVANTAGES OF PYTHON TECHNOLOGY IN FINANCIAL DATA ANALYSIS

A. Readability and consistency Software science and technology's ongoing development can facilitate the continuous absorption of Python technology's advantages, focusing on analyzing big data information's readability and consistency. There are newer, more efficient features than in other software applications.

The source code can run the data when Python technology is used. The data can be synchronously converted into a computer language that the Internet can understand by converting the data's source code into bytecode that computers can recognize. Python technology has become a highly consistent programming model and has achieved significant financial data analysis effects due to its relatively high readability and simple syntax.

B. Possess a high degree of cooperation and portability. The majority of current language program software possesses independent operation characteristics. The user can directly perform data operations on other programs and platforms, but they cannot alter any program data. This has been significantly altered by the Python program, which prevents the data operations performed by the program from being transferred to other platforms for continued operation. In addition, the Python program can provide numerous system interfaces and independent applets.

(MIJ) 2022, Vol. No. 8, Jan-Dec **e-ISSN: 2454-924X; p-ISSN: 2454-8103**

APPLICATION OF PYTHON TO FINANCIAL DATA ANALYSIS

A. Identify the crawler's search strategy. Web crawlers typically use a specific URL as their entry point, and the appropriate search strategy is identified by summarizing the URL's arrangement rules. Developers can discover the arrangement rules of the elements to be obtained through URL analysis. We talked about how the website pages with specific page numbers are distributed among the results obtained by searching for keywords. Usually, it is necessary to simulate the frequency with which real people visit the website to avoid drawing the special attention of network administrators and reduce the load on the server. As a result, when setting the frequency of retrievals, it is frequently necessary to first parse and store the information on the page before retrieving it after a three-second pause.

B. Get the HTML page's search information. In this step, developers must parse and analyze the HTML page's source code by calling the request library's get (URL) method. The page's content can be found in the stored Response object. Most of the core code is:

r = requests. get(url, header, timeout= 30)

return r. text

C. Remove valuable key information from HTML documents. After the HTML page retrieval information has been retrieved, removing the relevant, useful key information from the HTML document is necessary. You must be familiar with the HTTP protocol and HTML documents during this procedure to obtain crucial financial information like transaction volume, profit, and expenditure. The corresponding data list needs to be created during this procedure. The information about the business can be added to the goods list using the append method in the for loop.

D. Complete data collection and analysis. After parsing the HTML page, the developer can get the listed stock company's relevant financial status information. The data must be re-encoded to avoid messy codes appearing during the data analysis. As a result, different data types can be incorporated into data texts in various formats like CSV and XML, which is advantageous for subsequent processing.

PYTHON-BASED RISK QUANTITATIVE ANALYSIS

A. Variance method the description of asset returns can be used to determine the risk measurement. Markowitz developed the mean-variance investment portfolio theory in 1952. This was also the first way to measure asset risk. The equation is:

B. Downward Deviation Metric Roy also published the downside risk measurement method for measuring financial data risk in 1952.

The expected value of the asset, also known as the target rate of return, is the most crucial data in the calculation method for the following kind of volatility. Because his primary concern is the direction of asset returns, this method avoids the issues with the variance method. The formula is as follows:

$$\sigma(\mathrm{R},\ \mathrm{MARR}) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(min\,(R - MARR, 0))^2}$$

The minimum rate of return (MARR) is one of them. As the MARR, we frequently employ 0 or the average of the risk-free rate of return and asset return. Similarly, if the data exceeds the minimum rate of return, they will be considered zero in the calculation. On the other hand, if the data are less than the minimum rate of return, they will typically be factored into the formula for calculating the data. The obtained data can thus reflect the asset's downside risk without considering any upward fluctuations.

C. Value at Risk: The Python financial data risk measurement method underwent another change in 1994. JP Morgan Chase, a major investment bank, said that the value-at-risk method measures the risk of financial data. The formula is as follows:

$$P\gamma(X_t < \mathrm{Var}(\alpha, \Delta t)) = \alpha\%$$

Var is the anticipated maximum asset loss when the confidence and shortness target is met. In contrast, P represents the development of the event, or probability, while X in the formula represents the momentum of the asset portfolio and the value of financial assets within a particular time t. Using this formula, we need to know the probability of asset return to determine the value of financial data risk. To improve the accuracy of the data, we derive a probability distribution simulation method from this.

1) The Monte Carlo technique A statistical test or random sampling method are other names for this method.

It is a method for random simulation, as indicated by its name.

A suitable selection of parameters can increase the absolute data's reliability and accuracy through the reasonable and accurate construction of the data model. The method can also accurately reflect the negative and non-linear rate of return distribution at the same time. However, the results of subsequent calculations will be more affected by the random sampling data because it is somewhat dependent and may produce some incorrect data. Additionally, this method requires significant data and time to calculate. When the simulation is too large, this method should not be used.

2) The Covariance Matrix Approach The asset portfolio's hypothetical risk is estimated using this strategy.

The rate of return is naturally a normal distribution if we assume that all asset returns are positively volatile and that all data information has a linear posture. Then, to determine the distribution of the investment portfolio's return, we use the variance above method, the return on assets means, and other methods to estimate and calculate the investment portfolio's variance and mean.

CONCLUSION

The financial market is expanding daily, and fierce economic competition is a social development trend that cannot be avoided. Which parties that have benefited from big data will be able to lead the financial market and occupy the highest priority position in the big data society?

The best approach is financial data analysis in Python. Master the technology, evaluate its benefits and features and continuously adapt it to the company's growth. Opportunities and threats coexist simultaneously to ensure enterprise big data analysis's smooth operation. To ensure the sound and long-term growth of the financial data market, we must not only master Python technology but also perform well in risk measurement.

REFERENCES

1. Jingsong, & Wu. Financial Risk Quantification Research Based on Monte Carlo Simulation.
2. Li, C., & Liu, Q.. Prediction of Enterprise Credit Risk Based on Python Machine Learning.
3. Arias-Serna, M. A., Caro-Lopera, F. J., Echeverri-Arias, J. A., DA Castaneda-Palacio, & Murillo-Gomez, J. G.. (2017). Information system for the quantification of financial risk. Information Systems & Technologies (pp.1-6). IEEE.
4. Zhang, Y., Zheng, C., Zhang, Y., Zheng, M., & Wang, S.. (2019). Big Data Analysis of Financial Risk and Avoidance Strategy—Based on Futures Industry. 2019 6th International Conference on Behavioral, Economic and Socio-Cultural Computing (BESC).
5. Y An, & Shen, C.. (2017). Design and implementation of internet financial transaction risk identification system based on big data analysis.CECA, 42(2), 767-771.

6. (2019). Quantitative Strategy Design and Application Optimization Based on Python Language. Modern Computer.
7. Qi, J., & Yi, L.. (2016). Network financial fraud risk assessment system based on big data analysis. Journal of Computational & Theoretical Nanoscience, 13(12), 9335-9339.
8. Wang, S., Wang, H., & Economics, S. O.. (2019). Regional Financial Risk, Risk Exposure Dimension and Risk Prevention System: Based on the Data Analysis of Shanxi. On Economic Problems.
9. Yan, Hou, Ziyan, & Yuan. (2019). Financial risk analysis and early warning research based on data mining technology. Journal of Physics: Conference Series, 1187(5), 52106-52106.
10. Forst, T. M., & Murphy, P.. Development of a Python-Based Algorithm for the Quantification and Analysis of Multivariate Next Generation Sequencing Data Following Genome-Wide Mutagenesis and Cell Survival Screening.
11. Sun, A. Y., Jeong, H., A González-Nicolás, & Templeton, T. C.. (2018). Metamodeling-based approach for risk assessment and cost estimation: application to geological carbon sequestration planning. Computers & Geosciences, 113(APR.), 70-80.
12. Li, & Zhi-yuan. (2014). Grading prediction of enterprise financial crisis based on nonlinear programming evaluation: a case study of Chinese transportation industry. Abstract and Applied Analysis, 2014, 1-7.
13. Teng, H. W. , Lee, M. , & Lee, C. . (2019). Estimation procedures of using five alternative machine learning methods for predicting credit card default. Review of Pacific Basin Financial Markets and Policies.
14. Li, & Zhi-yuan. (2014). Grading prediction of enterprise financial crisis based on nonlinear programming evaluation: a case study of Chinese transportation industry. Abstract and Applied Analysis, 2014, 1-7.