

Technical Features and Trends of Data Science in Financial Engineering

Songyuanyi Lu

Linbinyuan Building B Room 402, Binhu New District, Hefei, 230000, China

Abstract: In the new financial era, the huge amount of data brings more challenges to the traditional financial business and creates unprecedented opportunities at the same time. In the financial industry, the use of data science by financial institutions has significantly deepened, from the traditional "data visualization presentation" to "data-based decision analysis". This paper analyzes the technical characteristics and development trend of data science in financial engineering against the background of rapid development of financial technology.

Keywords: Data science, Financial engineering, Technical characteristics, Development trend.

1. Introduction

Financial engineering is a cross-cutting discipline that deeply integrates theories and methods from disciplines such as finance, mathematics, and information and computing sciences. Data science requires mathematical statistics, computers, and specialized knowledge, etc. The process involves collecting, cleaning, transforming, and manipulating data, exploratory analysis, and modeling data using methods such as machine learning, and visualizing data as needed.[1] Quantitative investment in financial engineering is precisely the realization of the application of data science in the field of investment, and data science is important for the robustness of quantitative investment.

2. The Convergence of Data Science and Financial Engineering

In 2019, WEF released the report 《Data Science in the New Economy: A new race for talent in the Fourth Industrial Revolution》 [2]. The report argues that data science will be a fundamental skill in the future for industries such as artificial intelligence, machine learning, infocommunications, media and entertainment, finance and services, and is particularly important for economic growth. It is common to refer to information data as the "new oil" of the global economy, and data scientists are the people who can extract, refine and deploy this strategically valuable energy in the global economy.

Starting from the essential connotation of fintech, combined with the history and reality of fintech development, a basic analytical framework can be constructed about the evolutionary logic of fintech development. [3]The traditional financial service field cannot be separated from two main subjects, one of which is the customer and the other is the financial institution. The former is the object and demand side of financial services, while the latter is the supply side of financial services. Financial technology has injected two subjects with new connotation into the traditional financial service field: data and technology. The history of financial development is accompanied by the increasingly close integration of technology and finance. Technology has reshaped the financial industry, and the application of technology has given the modern financial industry a stronger

vitality.

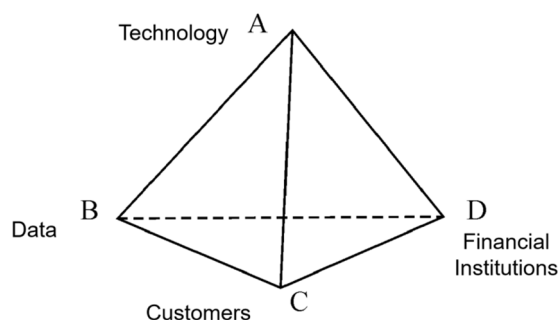


Figure 1. A framework for the integration of financial engineering and data science

3. Technical Characteristics of Data Science in The Field of Financial Engineering

3.1. Deep Integration of Data Model and Financial Business

Data science, on the one hand, focuses on the process of collecting, organizing and analyzing data, and on the other hand, data science itself is not limited to a specialized field, and any process of getting scientific knowledge based on data analysis is data science. As you can see, data science itself is applicable to different problems. Investment is a process of scientific decision making based on the data of various products in the financial market, which naturally conforms to the process of data science. [4]There is a large amount of data in the financial markets and some of it is relatively noisy, so data processing is an important element. In investment practice, quantitative investment has become the mainstream in the field of investment by collecting and organizing various kinds of data with noise from different channels, using algorithms to clean the data, and using machine learning algorithms to quickly calculate market laws and find investment opportunities.

There are many theoretical models in financial investment to help predict and optimize, yet there are few strategies that are consistently and consistently profitable in the market, mainly because the market has its own intrinsic laws, which

are usually understood only by experience, and those investors who have rich experience are not data scientists. At the same time, investors who use modern techniques such as quantitative often do not have enough investment experience, lacking investment experience on the one hand and understanding of data on the other. Data science emphasizes data, algorithms, and specialized domain knowledge, in the sense that it achieves a true convergence of modeling, programming, and financial engineering.[5]

3.2. Data-based Robustness and Credibility

In machine learning models for data science, robustness tests in financial measures are analogous to model tuning problems such as parameter selection, and in-sample out-of-sample tests in financial measures are analogous to resampling problems. James et al. in the classic Introduction to Statistical Learning: R-based Applications specifically address linear model selection and regularization, including compressed estimates such as ridge regression and LASSO, and use cross-validation and self-help methods to introduce

the resampling problem. This is exactly what robustness requires and extends it to more general levels such as over-fitting and under-fitting. In addition, Schutt and O'Neil also discuss causality in data science. The development of data science and the requirements of financial measurement practice make data robustness and trustworthiness even more prominent technical features of data science.[6]

4. The Development Trend of Data Science in The Field of Financial Engineering

The history of financial development is accompanied by the increasingly close integration of technology and finance. Technology has reshaped the business model of finance, and the application of data science has given modern finance a stronger vitality. Combining the history and practice of financial development, the historical evolution of the integration of finance and science and technology can be divided into three stages.[7]

Table 1. Historical evolution of the integration of finance and technology

Development Stages	Time Periods	Key Technologies	Key Features
Fintech 1.0	1866-1967	Telegraph, telephone and other analog information technology	A global telecommunication system built around transatlantic cables
Fintech 2.0	1967-2008	Computer Internet ATM machines	<ul style="list-style-type: none"> * Financial electronicization and informatization * Technology and finance are not fully integrated, and most technology companies mainly provide technology and data services or solutions for financial institutions * Banks' deposit and loan business is the main business * Finance is gradually transformed from an analog industry to a digital industry * The main business is dominated by traditional financial institutions * The mid to late 1990s entered the stage of Internet finance
Fintech 3.0	2009-present	Big Data, Cloud Computing, artificial intelligence and blockchain, etc.	<ul style="list-style-type: none"> * Comprehensive changes in the financial industry from business, business and operation models * Fintech startups are growing rapidly, traditional financial institutions are losing their dominance, and the two are moving from competition to cooperation, and the boundaries of financial services are becoming increasingly blurred

4.1. Fintech1.0 Stage

The emergence of Fintech can be traced back to the first half of the 19th century, a period of critical change in communications technology. Morse invented the telegraph in 1838, and in the following decades, electricity gained rapid development and telegraph transmission technology became increasingly sophisticated, with the first transatlantic cable successfully built in 1866. 1876 saw the invention of the telephone by Bell, and the first transatlantic telephone cable opened in 1956. These technological innovations laid the foundation for financial globalization. In today's Internet of Things era, it is hard to imagine a situation where the flow of information cannot be transmitted smoothly.[8]

The technological tools of this phase are analog information technology such as telegraph and telephone. The institutions and customers involved are basically larger financial institutions with a large business radius, conducting cross-regional operations and maintaining close business ties

with institutions outside the region. Technological development enables both the supply and demand of financial services to meet their immediate and cross-regional requirements to a greater extent. Data resources at this stage are initiated and provided by customers and financial institutions, and data are not characterized by monopoly or appropriation. At this time, data has not yet completed the logical transformation from instrumentalization to resourceization, and is still only a means of operation rather than an operational resource.

4.2. Fintech 2.0 Stage

Compared with the previous stage, the four major players in the development of Fintech have all changed. The technological means have developed into computer internet, ATM machines, etc., and the content of financial services has become richer and the cost has been further reduced. The scope of customers has expanded, and small and medium-sized customers can also obtain diversified financial services,

and their financial needs are not limited to important commercial or transactional needs, but can also be the financial needs of daily life. The meaning of data itself has also changed, and financial institutions have accumulated rich internal data, mined and formed structured data in the face of a large customer base. At this stage, it is mainly the financial institutions that make technology investments and initiate technology applications, and customers play more of a technology enjoyer role. In a certain sense, customers also have input into technology. For example, customers may pay a small fee for opening a financial account or credit card or for information alerts, but this is closer to the concept of a passive input and the amount of input is not comparable to the huge input from financial institutions. With technology, it is primarily up to the financial institution to use, invest in and master it. At the same time, financial institutions have some sense of ownership of the data. Financial institutions are willing to do so because technology improves both their own internal efficiency and their ability to serve their customers, thus enabling them to be more competitive in the industry. In addition to this basic motivation, financial institutions have the special benefit of turning data into an institutional monopoly of financial resources.

In Fintech 2.0, the nature of data has changed, both as a means and as a resource, completing the logical transformation from tool to resource, and in Fintech 1.0 and Fintech 2.0 before 2008, China was still in the period of awakening and initial development of Fintech. An important milestone in this period was the launch of Alibaba's Alipay in 2004. Alipay pioneered "secured transactions" and successfully opened the era of mobile payments in China.

4.3. Fintech3.0 Stage

In this stage, technology has made a major breakthrough, and new technologies such as big data, cloud computing, artificial intelligence and blockchain are developing rapidly. In addition to traditional financial institutions, other entities with large amounts of data, such as Alibaba, Tencent and other technology companies, have also penetrated into financial services. These companies generate a lot of data in other fields and are particularly capable of processing it. At the same time, the behavior of their own customers needs to be supported by financial services. For example, Alibaba launched Alipay for the important reason of meeting the financial needs of its customers and facilitating the transaction behavior between merchants and consumers. In this period, the meaning of financial institution has expanded, and whoever owns and can process data, while its customers have financial needs, it has the potential to enter the field as a new subject. In addition, the needs of customers and the composition of their groups have changed. Previously, they were financial services customers, who might need to make deposits, withdrawals, transfers, loans, or to buy stocks and bonds, public offerings, etc. At the same time, the meaning of data has also changed and become very complex. Most of the data information before is related to finance, but now other behavioral information of customers is also collected and recorded, and the structure of the data itself is more complex, and the scenarios of application are more abundant. The relationship between customers and data is also changing. From passive participation in providing data, customers are beginning to enjoy ownership of the value created by their own data under blockchain.[9]

Fintech 3.0 has grown rapidly in China, with 2008 as a

watershed year, when global financial regulation was generally tightened after the financial crisis. As traditional financial institutions struggled to cope with the strict regulatory framework, their business fronts shrank rapidly, leaving room for technology-based financial companies with fewer regulatory constraints to capture the market. The start-ups used big data and blockchain technology to provide a range of financial services such as payments, credit, insurance, and wealth management with better customer experience and higher efficiency, complementing a very large potential demand that traditional finance could not cover.

In this period, the concept of "Internet scenario-based" has broadened the channels and coverage of financial services and stimulated the financial consumption demand of a large number of potential users. The Internet financial model has formed a "full collection of transaction possibilities", breaking through the traditional boundaries of commercial viability. The development of mobile Internet technology has enabled the Internet, terminal devices and financial institutions to be effectively united, forming a new payment system that has changed the consumption habits of the whole society. At the same time, online lending, online financial insurance and other businesses are also developing rapidly.

The Fintech 3.0 stage can be further divided, and in 2013, Alipay successfully created the online money market fund BalanceBao, which was an important landmark in China's Fintech development. Since then, Fintech has reached new heights, with the injection of new technologies such as big data, cloud computing, Internet of Things, blockchain, financial cloud and artificial intelligence, and the introduction of face payment, smart investment advisor and smart customer service into the financial industry, renovating the financial industry at a rapid pace.

5. Conclusion

With the continuous combination of financial engineering and data science, a new industry has emerged in the investment field: smart investment advisors. Intelligent investment advisor (RoboAdvisor), which takes advantage of artificial intelligence, uses multiple algorithms and models to give investors comprehensive asset allocation services by combining their risk levels, expected returns, and market dynamics. The development of smart investment advisor relies on financial big data and algorithmic models, therefore, it is the most comprehensive application of data science in the field of financial investment, including financial data collection and processing, financial calculation and modeling, risk management, and visual communication with clients.

References

- [1] Huang Jiping. Quantitative Investment in the Era of Big Data: Functions, Challenges and Solution Paths [J]. Financial Journal, 2015(05):10-15.
- [2] Huang YuJie. Data science in the new economy: talent competition in the fourth industrial revolution [J]. Internet Economics, 2019(12):24-33.
- [3] Sun Jiadai, Bao Xingni, Zou Hanle. Innovation of Internet finance in the era of big data New Development and Risk Prevention and Control[J]. China Collective Economy, 2021(5): 91-92.
- [4] Yi Xianrong . The connotation, essence and future development of financial technology--a general analysis based on financial theory[J]. Journal of Jianghai, 2017, (02) : 13 - 20

- [5] [1]Ruilin Wang,Somayajulu Gowri Sripada,Nigel A Beacham. Auto-generating Textual Data Stories Using Data Science Pipelines[C]//.Conference proceedings of 2021 4th International Conference on Algorithms, Computing and Artificial Intelligence (ACAI 2021),2021:621-628.DOI:10.26914/c.cnkihy.2021.055297.
- [6] Ridgway Jim. Covid and data science: Understanding R0 could change your life[J]. Teaching Statistics,2021,43.
- [7] Hoepner Andreas G. F.,McMillan David,Vivian Andrew,Wese Simen Chardin. Significance, relevance and explainability in the machine learning age: an econometrics and financial data science perspective[J]. The European Journal of Finance, 2021, 27(1-2).
- [8] Jun Chen,Edward P K Tsang. Detecting Regime Change in Computational Finance:Data Science, Machine Learning and Algorithmic Trading[M].CRC Press:2020-11-02.
- [9] Chris Brooks,Andreas G. F. Hoepner,David McMillan,Andrew Vivian,Chardin Wese Simen. Financial data science: the birth of a new financial research paradigm complementing econometrics?[J]. The European Journal of Finance, 2019, 25 (17).