Proceedings of the 7th International Conference on Economic Management and Green Development
DOI: 10.54254/2754-1169/31/20231511

Study of the Applications of Alternative Data in the Field of Economics and Finance

Kangqiao Huang[1,a,*]

[1]*School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China*
a. huangciyo@buaa.edu.cn
**corresponding author*

Abstract: The term "alternative data" sometimes known as alternative data or non-traditional data, is most often used to describe regionally specific, valuable information that is distinct from typical financial data. Alternative data does not require a specialized data network or administration and may be collected and processed in real time. As a result, "alternative data" has quickly evolved in the financial sector in recent years. "Alternative data" is a well-known illustration of "big" data since, to start, it's enormously large in quantity as shown by its scope and transmission. The second element is the real-time or very close-to-real-time nature of the data collection and transmission. Data types and data structures come in a huge diversity, which is the third component. This paper studies the different classifications and acquisition methods of alternative data, focusing on the different applications of alternative data in the economic and financial fields, respectively, the application of alternative data in forecasting, especially in macro-economic forecasting and user income forecasting, and the application of alternative data in capital markets, business analysis and decision-making.

Keywords: alternative data, economics, finance

1. Introduction

Alternative data refers to data composed of non-traditional information or obtained through non-traditional means. Typically, enterprises in the economic and financial sectors use traditional data [1]. Compared with these traditional data, alternative data has advantages and broad application prospects in the financial and economic fields due to its large volume, diverse types, and fast flow speed [2].

This undoubtedly provides a new way for enterprises in the financial industry to expand. The application of alternative data can benefit from data analysis, risk management, and forward-looking forecasting [3].

The application of data substitutes for financial innovation is a powerful tool for breakthroughs and progress in the economic and financial fields. Nowadays, researchers are examining and researching the growth of alternative data in the field of economics and finance. Future developments and the current situation have significant theoretical and practical implications.

DOI: 10.54254/2754-1169/31/20231511

2. Classification and Acquisition of Alternative Data

2.1. The Notion of the Alternative Data

The term "alternative data" is used to describe non-traditional data, which is often used to describe data that is distinct from conventional financial data and has a specific use. Different from typical data is alternative data. It is not information collected by traditional means of information acquisition, information made public by exchanges, information received from insiders, or the most recent information made public by business announcements.

The number of Internet users is currently growing due to the Internet's growing popularity, particularly due to the widespread use of mobile smart terminal devices like smartphones and tablet computers, which encourages the ongoing improvement of the fundamental equipment functions for accumulating data usage. There is a chance that the huge volumes of data that couldn't previously be stored have now gathered and could be utilized successfully.

2.2. Alternate Data Classification

Nowadays, alternative data may be broadly categorized into three categories, primarily encompassing personal data, business process data, and sensor data [1]. Personal data is any information that is created as a result of a person's interactions with the outside world. Examples include product reviews, search histories, information people publish on social media, purchase histories, and shopping preferences. Business process data primarily consists of two types: business data gathered by alternative data businesses and data produced throughout the business process. Payment data, logistical data, and other types of data are among the data produced throughout company operations. The number of people using playgrounds, the number of people using big department shops, and other commercial statistics are all included in the data company's commercial data collection. Sensor data refers to data from GPS positioning, driving trajectories and alternative data created by personal devices, chemical plant data gathered by satellites through heat and light sensing, steel plant data, crude oil start-up, transportation and collecting data, etc.

2.3. Alternate Data Gathering and Collecting

The amount of substitute data is huge, the movement speed is fast, and there are many types, making it quite difficult to obtain. Currently, economic and financial institutions mostly use three technologies to collect alternative data: data can be obtained through the following methods: 1. purchasing directly from suppliers; 2 self-built teams; And 3 through the use of data collection tools. Among these three technologies, using data collection software has the greatest value and the lowest cost. Only a very small number of resources, including labor, can be used to collect alternative data. Therefore, when selecting data collection techniques, considering their practical value, some data collection tools and models are often used to collect data. For example, data mining and machine learning.

When using machine learning models to collect alternative data, commonly used machine learning models include logistic regression, support vector machines, adaboost [4], etc. Taking stock market forecasting as an example, recent literature data shows that machine learning models are generally superior to statistical and econometric models. With the shift from machine learning to deep learning, researchers can extract features and establish nonlinear correlation models without relying on econometric assumptions and human professional knowledge. In order to improve the accuracy of predictions, current research has been searching for new alternative data as the basis for predictions. Driven by the progress of natural language processing technology, the research on the synergy between quantitative online media and the stock market has exploded in the past few years.

DOI: 10.54254/2754-1169/31/20231511

The popularization of some data analysis technologies and the improvement of computing power have made it feasible to use more data for analysis.4. Application of alternative data.

3. Application of Alternative Data in the Economic and Financial Fields

3.1. Application of Alternative Data to Macroeconomic Forecasts

The application of alternative data contributes to the prediction of social macroeconomics. With the surge in available data, more and more alternative data such as social media data, large-scale GPS datasets, etc. can be used for predicting socio-economic phenomena. Moreover, alternative data has timeliness, and compared to official data with delay, alternative data prediction can obtain more timely and accurate results. Varian and Choi's groundbreaking article demonstrated the usefulness of Google search data for measuring various macroeconomic variables [5], while Giannone et al. proposed using unconventional data sources for high-frequency proximity prediction of US GDP growth [6]. Pappalardo et al. used big data analysis to make real-time predictions on the mobile data of French citizens [7]. Alternative data is gradually applied to the field of macro-economic forecasting.

3.1.1. Alternative Data used to Create CPI

With the rapid development of the Internet, online shopping has become very popular and has become the most popular shopping method for many people. Based on this, you can choose to capture price data online to collect and create a daily consumer price index. Taking the Indonesian Bureau of Statistics as an example [8], in order to generate daily CPI by capturing price data online, it is necessary to first select the retailer with the best credibility as the data collection object, then use network scraping technology to monitor the online prices of daily commodities, after collecting the data, run automatic programs to structure and clean up the data, and finally perform calculations and analysis to complete the creation of daily CPI, Visualization technology is used to display the results. Using this type of data to create CPI is time-sensitive and particularly convenient for creating daily CPI. And the cost of capturing price data from the network is lower compared to crowdsourcing mobile applications.

3.1.2. Alternative Data Used to Predict GDP

There are many types of alternative data, some of which can be used to predict GDP. Sandeep Kumar and his team propose that a country's carbon emissions data can be used to estimate GDP [9]. Especially for countries affected by war and inaccessible countries, using such alternative data to predict GDP is very effective.

3.2. Application of Alternative Data to Capital Markets

Alternative data such as social media, news reports, and weather data can be used for stock price forecasting, digital currency forecasting, and commodity price forecasting in the capital market.

In terms of stock price prediction, alternative data can be used to analyze factors such as a company's reputation, brand value, attention on social media, and news reports, thereby predicting the company's future performance and stock price fluctuations. For example, user reviews and feedback on social media can reflect consumer attitudes and perceptions of the company, thereby predicting the future performance of the company.

In terms of digital currency forecasting, alternative data can be used to analyze the market situation of digital currencies, discussions on social media, news reports, and other factors, thereby predicting the price fluctuations of digital currencies. For example, user reviews and discussions on

DOI: 10.54254/2754-1169/31/20231511

social media can reflect the market situation of digital currencies and the views of investors, thereby predicting the future price of digital currencies.

In terms of commodity price forecasting, alternative data can be used to analyze factors such as weather, natural disasters, and political factors to predict commodity price fluctuations. For example, weather data can be used to predict the production and prices of agricultural products, and political factors can be used to predict the price fluctuations of commodities such as oil and metals.

3.3. Application of Alternative Data to Business Analysis and Decision-making

Business analysis and decision-making are important components of enterprise management, and alternative data can provide strong support for these areas. The following are some examples of the application of alternative data to business analysis and decision-making.

3.3.1. Social Media Data Analysis

Enterprises can use social media monitoring tools to collect consumer feedback on their brands, products, or services. These data can help companies understand consumer preferences, needs, and attitudes, thereby optimizing product design and marketing strategies.

3.3.2. Sensor Data Analysis

Enterprises can use sensor technology to monitor data on production processes, logistics, transportation, equipment maintenance, and other aspects. These data can help enterprises optimize production processes, reduce costs, and improve efficiency.

3.3.3. Text Data Analysis

Enterprises can use text and table analysis tools to analyze text data. On the one hand, customer messages, comments, and emails can help companies understand consumer needs and opinions, thereby improving products and services. On the other hand, enterprises can establish a linear system model based on text table data, such as quarterly profit reports, to forecast the daily income of the company [10].

3.3.4. Audio and Video Data Analysis

Enterprises can use audio and video analysis technology to analyze customer service calls, conference records, and other data. These data can help enterprises understand customer needs and service quality, thereby improving customer service.

3.3.5. Network Data Analysis

Enterprises can use network analysis technology to understand competitors, market trends, and industry development trends. These data can help enterprises formulate better marketing strategies and business development plans.

In short, alternative data can provide enterprises with more comprehensive, accurate, and real-time information, helping them become more accurate and efficient in business analysis and decision-making.

3.4. Application of Alternative Data to Forecast User Revenue

Alternative data can be applied to enhance user revenue estimation models. In a Columbia University study, the advantages of alternative data for super applications in enhancing user revenue

DOI: 10.54254/2754-1169/31/20231511

estimation models were explored [11]. The main applications of alternative data include personal information, consumption patterns, payment information, and financial services. A Super-App alternative data model was established to compare the performance of alternative data sources with the performance of locally recognized revenue estimators that only consider financial system information. Research shows that alternative data can obtain information that is not available in local revenue estimators. Alternative data has the ability to predict revenue estimates.

3.5. Application of Alternative Data to Socio-economic Analysis

Some alternative data can be applied to socio-economic analysis, such as administrative data such as taxes and hospital records, business data such as credit card transactions, and various text data in online searches and social media. These data can be applied after being extracted and organized. For example, during the COVID-19 pandemic, these unconventional alternative data were used to supplement official statistics and played a significant role. The application and development of alternative data as powerful decision-making tools can help decision-makers analyze and design policies conducive to socio-economic growth and social welfare.

4. Conclusion

This paper mainly introduces the classification and acquisition of alternative data, focusing on the different applications of alternative data in the economic and financial fields. This paper mainly analyzes from four aspects, and respectively studies the specific applications of alternative data in macro-economic forecasting, capital market, business analysis and decision-making, and user income forecasting.

Due to the characteristics of timeliness and diversity of alternative data, which make up for the shortcomings of traditional financial data, it has broad applications and promising development prospects in the fields of finance and economy. But there are also some shortcomings. Firstly, there may be issues with the quality and accuracy of alternative data, and data cleaning and processing are necessary to obtain reliable analysis results. Secondly, due to the diversity and complexity of alternative data, more manpower and resources need to be invested in analyzing these data. At the same time, this paper also finds that the current application of alternative data in the financial and economic fields is mostly text data, while the application of image and video data is still limited. In the future, more application forms of alternative data can be explored, making alternative data more diverse in applications. We should be innovative and open to new ideas while gathering and utilizing "alternative data". The fundamental principle of "alternative data" is to break through conventional barriers, utilize all of the many sorts of data now accessible, and use a sword to solve issues.

In addition, the privacy and security issues of alternative data also need to be taken seriously. For example, when collecting user comments and behavior data on social media, it is necessary to comply with relevant laws and regulations to protect user privacy. But the society also needs to achieve the openness and sharing of alternative data across platforms to improve the application efficiency of alternative data in the financial and economic fields.

In summary, although alternative data has certain applications in the financial field, it still needs to face many challenges and shortcomings, and these problems need to be better addressed in order to better play its role.

References

[1] *Liu, Y.: The development status and future trends of alternative data in the financial field. China High and New Technology (19), 104-105+130 (2021). (in Chinese).*

DOI: 10.54254/2754-1169/31/20231511

[2] *Liao, L.: Alternative data: New highlights of economic growth. People's Forum (Academic Frontiers) (6), 22-27 (2021). (in Chinese).*
[3] *Xie, L., Tang, M., Ren, B.: Research on the application of "alternative data" in the field of technology financial services. China Science and Technology Resources Guide 51(05), 15-20 (2019). (in Chinese).*
[4] *Dierckx, T., Davis, J., Schoutens, W.: Using machine learning and alternative data to predict movements in market risk. arXiv: Computational Finance, (2020).*
[5] *Varian, H., Choi, H.: Predicting the present with Google trends. Economic Record 88(5), 2-9 (2009).*
[6] *Giannone, D., Lenza, M., Primiceri, G.E. Economic predictions with big Data: The illusion of sparsity. Liberity Street Economics, Federal Reserve Bank of New York, (2018).*
[7] *Pappalardo, L., Vanhoof, M., Gabrielli, L., et al.: An analytical framework to nowcast wellbeing using mobile phone data. International Journal of Data Science and Analytics 2(1), 75-92 (2016).*
[8] *Manik, D.P.: A strategy to create daily consumer price index by using big data in statistics Indonesia. In: International Conference on Information Technology Systems & Innovation. IEEE, (2016).*
[9] *Kumar, S., Shukla, A.K., Muhuri, P. K., et al.: Transfer Learning based GDP Prediction from uncertain carbon emission data. In: 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). IEEE, (2019).*
[10] *Fleder, M., Shah, D.: Forecasting with alternative data. In: SIGMETRICS '20: ACM SIGMETRICS / International Conference on Measurement and Modeling of Computer Systems. ACM, (2019).*
[11] *Suarez, G., Raful, J., Luque, M.A., et al.: Enhancing user's income estimation with super-app alternative data. arXiv, 05831 (2021).*