Financial Analytics With Big Data

F

Leon Wang
Big Data Club, USA

INTRODUCTION

Finance is a fundamental part to everyone's life. As the big data, Internet of Things, cloud computing, and other ideas and technologies are integrated into social life, the big data technology can improve the corporate financial data processing. At the same time, with the fiercer competition between enterprises, investors and enterprises have paid more attention to the role of financial crisis warning in corporate management. The development of modern information technologies entails an unprecedented growth in the volume of computing resources and large data sets. Thanks to this, ML methods available through open-source toolkits are gaining popularity among analysts and developers. Deep learning models have proven extremely successful in a wide range of applications, including image processing, learning gamification, neuroscience, energy conservation, and medical diagnostics (Shang, Lu, & Zhou, 2021; Zhou, Lou, & Jiang, 2019).

The finance industry has adopted ML to varying degrees of sophistication. Several mathematical disciplines, including statistical computing, data mining, financial econometrics, probabilistic and dynamic programming were brought together by ML in finance. There are many misconceptions and limited understanding of the possibilities of this field. Effective ML methods remain poorly understood and often mathematically unsound (Sreejanya, Chowdary, Atri, Reddy, Harith, & Mahajan, 2022). A key challenge to understanding ML is the lack of well-established theories and concepts that are necessary for financial time series analysis. The key to implementing ML in finance is to be able to run ML alongside parametric methods, observing over time the differences and limitations of parametric modeling based on fit metrics in the sample. Karachun, Vinnichek, & Tuskov (2021) address many finance practitioner's concerns that neural networks are a "black-box" by showing how they are related to existing well-established techniques such as linear regression, logistic regression, and autoregressive time series models. Neural networks can be shown to reduce to other well-known statistical techniques and are adaptable to time series data. Statistical tests must be used to characterize the data and select the algorithm. If the data is of sufficiently high quality and adds a new source of information, then it can be easily scaled up. ML is a more reliable approach than many methods of parametric financial econometrics used today. At the same time, the use of ML requires strong skills of scientific justification. And it is not a panacea for automatic decision-making.

BACKGROUND

There has been an extremely fast growth in the amount of data that is made and collected on a daily basis and because of this data collection, analysis, and processing have created opportunities for new technology, jobs and industries. According to Forrester (2019), big data is considered as a vital dominant driver of competitive advantage that refers the ability to outperform the rivals for businesses. Today, there is

DOI: 10.4018/978-1-7998-9220-5.ch114

a dramatic increase in the amount of generated, mined and stored data, reaching a market size of $50 billion to reach $104.3 billion by 2026 (Basdas & Esen, 2021). Companies produce large amounts of raw data in daily basis via IoT, smart devices and cloud platforms. Technology is constantly evolving and changing the way that we do business. The implementation of Big Data in the business world has proved to be very beneficial and effective especially in the financial sector. The addition of this new research will allow companies to combat fraud, offer more personalized experiences to their customers, and make smarter investment decisions.

Financial services industry includes SME finance, wealth and asset management, mobile payment platforms and the popular cryptocurrency. Due to the intricacy and sheer magnitude of the data these platforms create, they are also prone to many arising issues which is why financial analysts are highly sought after. These are especially important to analyze because big data and internet finance are being regarded by more and more scholars to be the "new engine of growth" (Wang, 2018). Big data has already proved its importance and value" (Ke & Shi, 2014). Identifying the opportunities, challenges and implications of big data in finance help to further highlight areas of growth for organizations. With ever evolving technology and with different types of data now continuously available with the advancements of information technologies, data has become, one of if not the most, valuable commodity in the financial services industry. Finance is at the core of modern economic operations, and financial viability determines the quality and potential of the overall economy (Wang, 2018). Most areas of business and finance are heavily connected with big data (Hasan, Popp, & Olah, 2020).

Basdas and Esen (2021) provide a systematic review on the world's revolution in information and communication technologies (ICTs) over last couple of decades. Big data appeared as a revolutionary phenomenon that influenced decision-making processes. The need for big data towards the digitalization of services, utilization of social media and new channels to reach customers, demand for personalized services and continuous flow of vast amount of data in the sector is growing exponentially. Massively parallel processors and modern data management architectures have led to more efficient operations and a better decision making for companies to process and analyze such complex and large-scale data. Especially, financial services companies leverage big data to transform their business processes. Heuristic decision making, simple reporting and statistical analysis dominated in the 1960s and 1970s. Relational or hierarchical databases in table based format were organized to store data in the 1990s. Extract, transform and load (ETL) processes have been implemented for cross-functional activities, fast query processing and multiuser environment in order to help enterprise data mitigate from day-to-day transactions to data warehouses. Companies started to focus on value creation by operational data warehouses that accumulate business transactions in the early 2000s. Companies became more analytical and data-driven in the following decade due to their ability in processing high volume and velocity of data to realize business benefits. The advent of wireless has avalanched the variety of data and brought a proprietary resource for companies in 2010. Virtual infrastructures, a real-time communication network, allowed many companies to allocate their resources for moving, storing and analyzing the huge amount of data over the second half of the 2010s. Further developments in the digital transformation enable companies to move with big data solutions using Artificial Intelligence (AI) for organizing and performing business tasks today.