AI

MDPI

Review

Data Science in Finance: Challenges and Opportunities

Xianrong Zheng [1,*], Elizabeth Gildea [2], Sheng Chai [3], Tongxiao Zhang [4] and Shuxi Wang [5]

[1] Information Technology & Decision Sciences Department, Old Dominion University, Norfolk, VA 23529, USA
[2] School of Cybersecurity, Old Dominion University, Norfolk, VA 23529, USA; egild002@odu.edu
[3] School of Computer Science and Information Systems, Northwest Missouri State University, Maryville, MO 64468, USA; schai@nwmissouri.edu
[4] School of Computer and Communication Engineering, Northeastern University at Qinghuangdao, Qinghuangdao 066004, China; zhangtongxiao2@sina.com
[5] Department of Artificial Intelligence, University of International Business and Economics, Beijing 100029, China; wangshuxi@uibe.edu.cn
* Correspondence: x1zheng@odu.edu

Abstract: Data science has become increasingly popular due to emerging technologies, including generative AI, big data, deep learning, etc. It can provide insights from data that are hard to determine from a human perspective. Data science in finance helps to provide more personal and safer experiences for customers and develop cutting-edge solutions for a company. This paper surveys the challenges and opportunities in applying data science to finance. It provides a state-of-the-art review of financial technologies, algorithmic trading, and fraud detection. Also, the paper identifies two research topics. One is how to use generative AI in algorithmic trading. The other is how to apply it to fraud detection. Last but not least, the paper discusses the challenges posed by generative AI, such as the ethical considerations, potential biases, and data security.

Keywords: data science; financial technologies; algorithmic trading; fraud detection

check for updates

Citation: Zheng, X.; Gildea, E.; Chai, S.; Zhang, T.; Wang, S. Data Science in Finance: Challenges and Opportunities. *AI* **2024**, *5*, 55–71. https://doi.org/10.3390/ai5010004

Academic Editor: Luis Javier Garcia Villalba

Received: 12 October 2023
Revised: 21 November 2023
Accepted: 2 December 2023
Published: 22 December 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Data science has played a major role in business, finance, healthcare, science, engineering, etc. [1,2]. In these fields, data can be studied, analyzed, and used to benefit society. Data science includes a wide range of new systems and techniques. However, these new techniques all bring forth their own opportunities and challenges. The main objective of this research is to describe how data science is used in finance and to understand the unique challenges and opportunities.

1.1. Data Science

Data science is an interdisciplinary field that draws from computer science, math, statistics, and more to provide new opportunities for research [3]. Longbing Cao, an AI and data science researcher, drafted the formula below to acknowledge the fields that provide insights into data science [4,5].

data science = {statistics ∩ informatics ∩ computing ∩ communication ∩ sociology ∩ management | data ∩ domain ∩ thinking}

where "|" means "conditional on".

It is important to first understand data science as well as other terminologies that are used in this paper, as shown in Table 1.

1.2. Data Science Tools

Data science uses many tools and techniques to turn data into meaningful information. These include machine learning tools, such as TensorFlow [6] and Natural Language Toolkit [7], programming tools, such as Python and R programming, data analysis tools, such as SAS and MATLAB, and big data tools, such as Apache Hadoop. As shown in Figure 1, there are many tools that aid in data science.

Table 1. Terminologies used in data science.

Term	Definition
Data Science	An interdisciplinary field that combines statistics, mathematics, and computing to extract insights from data sets
Big Data	Very large amounts of data being collected with a great variety of data types and a high rate of velocity
Machine Learning (ML)	Computers learning from data instead of being programmed
Artificial Intelligence (AI)	Machines using mathematical functions and complex algorithms to perform tasks that humans perform
k-means	A clustering method that groups items by similar traits
Neutral Networks	A type of deep learning that uses AI to teach computers to act like the human brain

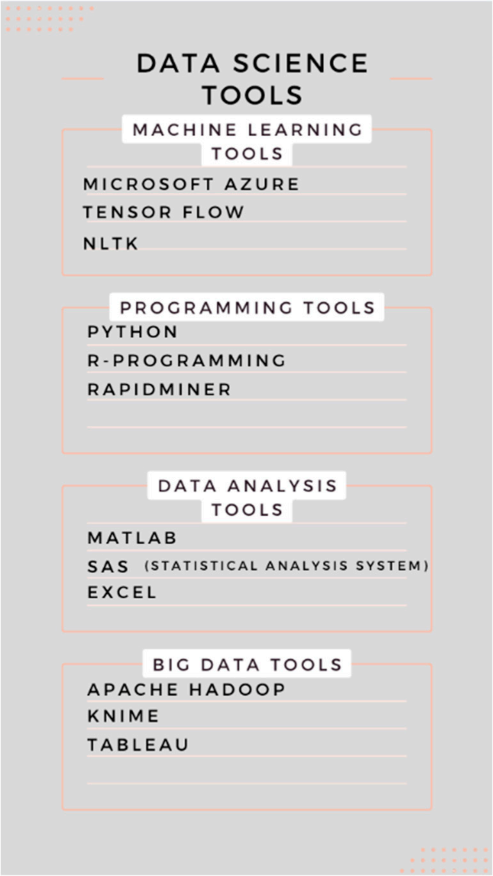

Figure 1. Data science tools.

With tools like these, data can be turned from numbers, facts, and statistics into meaningful information. These data science tools help to clean the data sets to clear out inconsistencies and repetitive data, transform the data to be more understandable, model the data to best fit business needs, and then help to visualize and communicate data to clients. Without these tools, the process of understanding data would be much more time consuming and complicated. Table 2 further describes some popular tools used in data science.

Table 2. Popular tools used in data science.

Tool	Description
TensorFlow	Open-source software that performs model evaluation, data processing, and visualization for machine learning. It is used by companies, such as Airbnb, Spotify, Qualcomm, and PayPal [6].
Natural Language Toolkit (NLTK)	Open-source library "with a suite of text processing libraries for classification, tokenization, stemming, tagging, parsing, and semantic reasoning, wrappers for industrial-strength NLP libraries, and an active discussion forum" [7].
Python	A dynamic and flexible programming language. Python includes numerous libraries, such as NumPy, Pandas, Matplotlib, for analyzing data quickly [8].
R Programming	An open-source programming language and environment for developing statistical computing and graphics [8].
SAS	A comprehensive tool suite, including visualizations and interactive dashboards, for analyzing, reporting, data mining, and predictive modeling [8].
MATLAB	A high-performance numeric computing platform that aids in the analysis and visualization of data.
Apache Hadoop	A framework that allows for the distributed processing of large data sets across clusters of computers, using simple programming models [9].

1.3. Data Science in Finance

Data scientists have used data to create meaningful insights, which has resulted in a rise in revenue in multiple disciplines. This paper will focus on how data science has developed in finance. Data science can allow for direct information that is diverse, reliable, and delivered fast. In the finance world, these data science tools will allow for more predictive power [10]. According to Pisoni et al. (2021) [11], "The information system architecture may include algorithms related to digital financial services, e.g., digital factoring, invoicing, and loan calculations. Furthermore, it may contain technologies that support digital investments, trading, crowdfunding, digital money, virtual currencies, and digital payments." This means that finance is shifting from traditional human traders and cash transactions to a virtual and safer environment. Data science has aided in the creation of FinTech, prevented fraud, and helped to develop robo-advising. Data science has provided the resources for finance to take off in a whole new light.

The opportunities that data science has provided for finance do not come without their challenges. Fraud protection, algorithmic trading, and robo-advising are unique outcomes that present improvements in the industry, but still produce negative outcomes. For example, FinTech has created fast and convenient online transactions; however, online fraud rose by 149% from Q4 of 2020 to Q1 of 2021 [12]. Algorithmic trading can be very beneficial and fast; however, can it catch all errors that human traders do?

The paper is organized as follows. In Section 2, we discuss FinTech, what it is, how data science is involved, how AI has been introduced, and the challenges and opportunities associated with FinTech. Section 3 focuses on algorithmic trading and how data science is involved with it. Section 4 focuses on financial fraud detection. Section 5 presents the conclusion.

2. Data Science for Financial Technology

Financial technology is the growing sector of automated banking, online payments, algorithmic trading, cryptocurrencies, etc. It essentially removes the cost of financial intermediation which, in 2016, averaged 2% of the transaction cost [13]. Most people use a FinTech service everyday, including Venmo, PayPal, Credit Karma, Robinhood, and Apple Pay. It is convenient for many people as well as for the big companies providing it. In the late 1960s, ATMs were invented and that was the start of FinTech. It took out the cashier or banker giving the customer money. Now, we make cashless transactions in the blink of an eye, without even thinking about carrying cash.

Technology companies are diving headfirst into the finance world. Apple, Google, and Alibaba all have their respective payment platforms: Apple Pay, Google Pay, and Alipay. Bill Gates stated back in 1994 that, "Banking is necessary, banks are not" [14]. These technology companies will most likely never go fully into the banking world, as there are too many regulations to follow, but they will still benefit from the partnerships that they have with big banks.

2.1. AI and Data Science in Financial Technology

FinTech is growing and so is the use of data science and AI in the field. Companies are gathering so much data than they are able to process and learn from. These data help them to stay ahead of competitors and grow in their market. Data science gives them the ability to collect, process, clean, analyze, model, and communicate the findings to be beneficial to their companies. AI helps to automate this process, make it fast, find trends that humans cannot, make employees more productive by eliminating some of their repetitive and time-consuming tasks, and help to protect the companies and clients.

The convivence of FinTech for the public will make Apple, Google, Alibaba, and other companies in the FinTech world successful. However, their success would not be attainable without the use of data science.

Data science in FinTech allows us to identify trends, make predictions, and gain insights. FinTech companies can leverage data science along with AI to implement successful changes in the companies. The two main technologies in FinTech that help to obtain meaningful information from data are artificial intelligence (AI) and blockchain. AI is changing how FinTech companies provide for their customers, and it gives them a competitive advantage over their competitors who do not use it. AI allows companies to analyze data in a fraction of the time that a human would take. Data science processes, analyzes, and makes predictions with data, and AI enhances how it is performed.

AI can make future predictions, identify fraud, and give more accurate advice to customers. For example, AI can help a small business navigate prices for their products to be sold at, identify whether cybercriminals are making fraudulent transactions, or help a customer who wants to learn how to invest their money. Below is information on how each of these things are performed.

Future Predictions Enabled by Artificial Intelligence

AI uses past historical and present data that have been processed and analyzed with data science tools to make future predictions. These predictions can involve future revenue, what price point that products should be sold at, or when a company should buy or sell stock. Past and present data that are being collected can be transformed in seconds into meaningful future predictions. A way that AI analyzes data is through clustering and classification. Clustering the data is a way of grouping similar data points, while classification involves labeling each smaller group. Clustering can be performed with popular tools, such as *k*-means and neutral networks. AI is also able to find patterns in data that a human would not be able to find. These patterns lead to very insightful comparisons between past and present data sets that lead to insightful future ideas.

It is important to note that there are issues with future predictions made with AI. While the data are real and accurate, AI machines cannot deal with sudden unexpected events that are not the norm, which then can lead to inaccurate predictions. Also, it can be hard to make business leaders trust the future predictions that AI machines create. Although it is based off their own companies' data, it is almost impossible to explain how it is making the predictions. The theories can be explained, but knowing what neuron in the AI 'brain' made the prediction is not possible. It will take many years for full trust of AI to occur, if ever.

Fraud Detection Enabled by Artificial Intelligence

As mentioned before, AI can identify patterns more accurately and faster than humans. Within these patterns, these machines can identify anomalies that may be fraud. It is also helpful that AI is constantly learning and can learn what is fraudulent and what is not.

There are many falsely flagged transactions that are thought to be fraud but are not. AI can learn and eliminate those falsely flagged fraudulent transactions. Artificial intelligence machines also help to monitor transactions at all hours. There is no longer a need for a human to monitor charges at 3 a.m. because the AI machine can instead. This increases the likelihood that fraud will be detected and responded to quickly.

To protect fraud on digital banking channels, which have become increasingly popular, AI uses biometrics, such as iris and fingerprint scanners, to verify the identity of the user. This helps to restrict access to sensitive information.

Some issues with AI in fraud protection is that there can be falsely flagged transactions. Although AI is learning, these are an inconvenience to companies and their clients. Again, it is hard to trust AI in these situations. Will it always work? Why is it marking this as fraud when it is not? These are questions that people might have, and it is hard to give exact answers.

Artificial Intelligence Helping Customers

AI can help customers in a variety of ways. AI chatbots can help customers online in a very productive way. They can answer questions just like a human would and can provide real-time help to customer issues. This is convenient because questions can be answered no matter the time of day, and the AI machines can be a supplement to the employees. They can also provide answers in multiple languages and will not become flustered due to stress.

AI can also increase customer engagement. AI can monitor what the customer clicks on and interacts with, and then they can provide more personable experiences for the customer. This results in happier customers and more sales for a company.

The issues with AI in this regard are that it cannot understand and sympathize with a customer. If a customer calls their bank in a frantic and upset tone saying that they have lost their credit card, the AI chatbot will go through the same process as always. Sometimes, humans are overwhelmed and stressed and need a human to respond with the appropriate level of sympathy.

In the Market Analysis Report, it was stated that, “Artificial Intelligence in FinTech market size was valued at USD 9.45 billion in 2021 and is expected to grow at a Compound Annual Growth Rate (CAGR) of 16.5% from 2022 to 2030” [15]. AI is no longer the technology of the future. It is here and it is here to stay. The growth of AI has been attributed to key contributors like Google, Intel, Microsoft, and Oracle, but it is trickling down into small companies and startups. AI will continue to transform the FinTech industry by creating future predictions, identifying fraud, and aiding in customer service.

How to apply generative AI to finance is a topic of interest. Generative AI has the potential to transform the finance industry in the coming years. It can automate tasks, improve decision-making processes, and enhance the overall efficiency in finance. For example, generative AI can analyze news articles, social media, and other textual data to sense market sentiments. Also, generative AI can power chatbots for customer service, answering questions, and providing information about financial products and services. It should be noted that generative AI poses some challenges, including ethical considerations and potential biases. It is crucial to deploy responsible generative AI in finance.

2.2. Blockchain in Financial Technology

Blockchain is a database that stores data in blocks with cryptography. Blockchain can provide help and opportunities for FinTech, such as making it more secure and faster, and it can aid with financial inclusion. Blockchain is decentralized, meaning that there is no one large group in charge. Instead, control is distributed on a network. This allows more transparency, as the public can view all transactions. There is also an immutable ledger that means the integrity of the data is maintained for security purposes. One of the main challenges for FinTech is security, and blockchain helps to improve the transparency and integrity of data, which has started to make FinTech more secure.

As mentioned, there is more transparency in blockchain, due to its decentralized nature. This transparency reduces the need for audits and intermediaries. Instead, there are peer-to-peer transactions (transferring money from one user's account to another user), which speeds up the transaction process. It is fast to transfer and receive money with them, which has become extremely popular in the past few years with apps, such as Venmo, PayPal, and Cash App.

Financial inclusion, which is mentioned later as a challenge associated with FinTech, is the idea that not everyone has access to the devices or Internet needed for financial technology. Although this is a large challenge, blockchain can allow people without access to a bank to transfer money. Money transfers can be performed 24/7 by anyone with a digital device and Internet access. Blockchain also allows for cheaper transaction fees, which can help those with a financial burden to transfer money.

2.3. Opportunities and Challenges of Financial Technology

FinTech is going to continue to provide digital payments, peer-to-peer transactions, and international payments. It is going to continue to provide financial services faster and more efficiently with the help of technology. As previously mentioned, data science is helping FinTech to create better future predictions, have better fraud detection, and provide better help and feedback to customers. The FinTech industry will continue to pursue its goal of make more improved and personalized experiences for the customer.

Financial technology, which has made some customers' lives so much easier, does not come without its unique set of challenges. FinTech creates issues with regulatory compliance, financial inclusion, and data privacy/security.

Regulatory Compliance Issues with Financial Technology

FinTech companies are developing and using the newest and best technology on the market. This technology is being delivered faster than frameworks for regulations are being developed. The concepts that a company may want to implement might not comply with existing regulations. Then, there is a gap in what the FinTech companies can provide and what the regulations allow them to provide. Getting the regulations updated and followed can be a time-consuming process. The regulations that a company must abide by also vary depending on the country. For example, in the United States, a FinTech company must abide by the Gramm–Leach–Bliley Act to ensure consumer data protection, while in Europe, it is the General Data Protection Regulation that must be followed. These regulations are a challenge to follow, and they are costly and time-consuming, but they are extremely important to abide by.

Financial Inclusion Issues with Financial Technology

FinTech can lead to certain populations being excluded from the benefits of the field. There are many issues with getting the technology needed for FinTech. "Financial exclusion could arise from various sources, including the lack of access to digital infrastructure, such as mobile phones, computers or the Internet, financial and digital illiteracy, potential biases in algorithms, and/or lack of trust" [16]. Technology and Internet access can be expensive and complicated to use. It is difficult for certain populations to understand how to use technology correctly. There is also potential for biases in algorithms. This means that the existing FinTech algorithms are not always fair and that "Algorithms used by banks to predict credit card debt defaults typically favor wealthier white applicants" [16]. These algorithms can discriminate based on race, gender, and income. FinTech can be exciting and helpful for many people, but it also excludes many people who do not have the resources for it.

Data Privacy/Security in Financial Technology

FinTech companies collect a lot of personal and financial information, such as names, addresses, social security numbers, bank account numbers, and personal biometrics that might be used to gain access to an account. All of this information is "gold" to hackers and makes these financial technology institutions vulnerable to ransomware, malware, phishing attacks, and data breaches.

A recent example is a cyberattack on the Industrial and Commercial Bank of China (ICBC), which is the world's largest lender by assets [17]. On 9 November 2023, the U.S.

financial services division of the ICBC experienced a ransomware attack that disrupted the trading of Treasuries. LockBit 3.0 was the software behind the attack. It can enter an organization when someone clicks on a malicious link in an email and then extract sensitive information about a company. LockBit 3.0 is evasive and hard to detect. It is a challenge for security researchers. LockBit is the most popular strain of ransomware, accounting for around 28% of all known ransomware attacks from July 2022 to June 2023. The business model of the LockBit group is "ransomware-as-a-service". It sells its malicious software to other hackers, who then carry out the cyberattacks. According to the U.S. Department of Justice, LockBit actors have launched over 1400 attacks in the United States and around the world, demanded over $100 million in ransom, and received tens of millions of dollars in payments via bitcoin.

It can be a challenge for these companies to operate safely and protect their customers. It is also difficult to keep data safe, when there is a plethora of technologies being used, such as different types of hardware and software, and these technologies are being used all around the globe. As stated before, every country has different regulations and policies, so it is difficult to have a strategy against cyberattacks in an uncoordinated environment. Data privacy and security concerns are also a reason why people do not want to partake in the utilization of FinTech. It can be hard to trust that your data will be safe in these environments.

To protect customers' data, it is important to have encryption, multifactor identification, and firewalls. Since data might be shared with third parties, it is important for text to be encrypted into ciphertext, so that it is unreadable during the transmission of data. Multifactor identification can help to protect data security and privacy, because it means that more than a password and username is needed to access data. This means that there is a more in-depth security process to access sensitive data. Lastly, all FinTech companies should have firewalls to monitor traffic coming in or out of their network. This can help them to block and identify suspicious activity.

3. Data Science for Algorithmic Trading

Another key outcome of data science in the FinTech sector is algorithmic trading. Algorithmic trading is a way to execute orders with pre-programmed algorithms to automate trades, depending on the price, time, and volume. Algo trading uses very complex mathematical calculations and rule-based algorithms to determine whether trades should be executed. Algo trading uses machine learning to understand the past trading history and patterns. The accuracy and efficiency of algorithmic trading depends heavily on the analysis of data sets to determine the correct and best practices. Throughout this section, the history of algo trading is described, as well as how data science is involved, and the trends, challenges, and opportunities associated with algo trading.

A trading process includes five main stages: data access/cleaning, pre-trade analysis, trading signal generation, trading execution, and post-trade analysis [18]. Developers use different approaches, including back-testing and optimization, to evaluate the effectiveness of the algorithms.

Data Access/Cleaning

In the trading process, it is necessary to obtain market data first, including the stock price, trading volume, and other financial data. Then, the data need to be cleaned and preprocessed to provide accurate data for the subsequent analysis and decision making.

Pre-trade Analysis

This includes the development of trading strategies and the setting of trading objectives, etc. Before actual trading, traders need to evaluate and study the market. By using various analytical tools and techniques, traders can identify potential trading opportunities and risks.

Trading Signal Generation

Based on the results of the pre-trade analysis, traders generate trading signals indicating when to buy or sell a specific financial asset, according to the established trading strategy. Also, the traders decide which financial instruments to invest in over the next time horizon.

Trading Execution

Once the trade signal has been generated, the traders execute the actual trade. This involves executing buy or sell orders and ensuring that trades are carried out based on established strategies and rules. Trade execution can be performed manually or with the help of automated trading systems.

Post-trade Analysis

The outcome of the trading activity, such as the difference between the expected price and the final strike price and the profit and loss statement, is assessed.

3.1. The History of Algorithmic Trading

The outline of the history of algorithmic trading is featured in Figure 2. What used to be conducted on the stock exchange floors with people yelling out trades is now being conducted by supercomputers that execute trades faster than a human could click. In 1969, Electronic Communication Networks (ECNs) were introduced to help match buyers and sellers, taking the middleman out of the equation. Overall ECNs also help to reduce trading errors and lower prices. In the 1970s and 1980s, the NASDAQ started to use computerized trading, leading it to be the first electronic stock market that allowed for less human interaction and more efficient trading. This made trading more available and faster. Algorithmic trading in the mid-1990s, "accounted for only 3 percent of the market, [compared] to recent times when it has reached almost 85% of dollar trade volumes" [19].

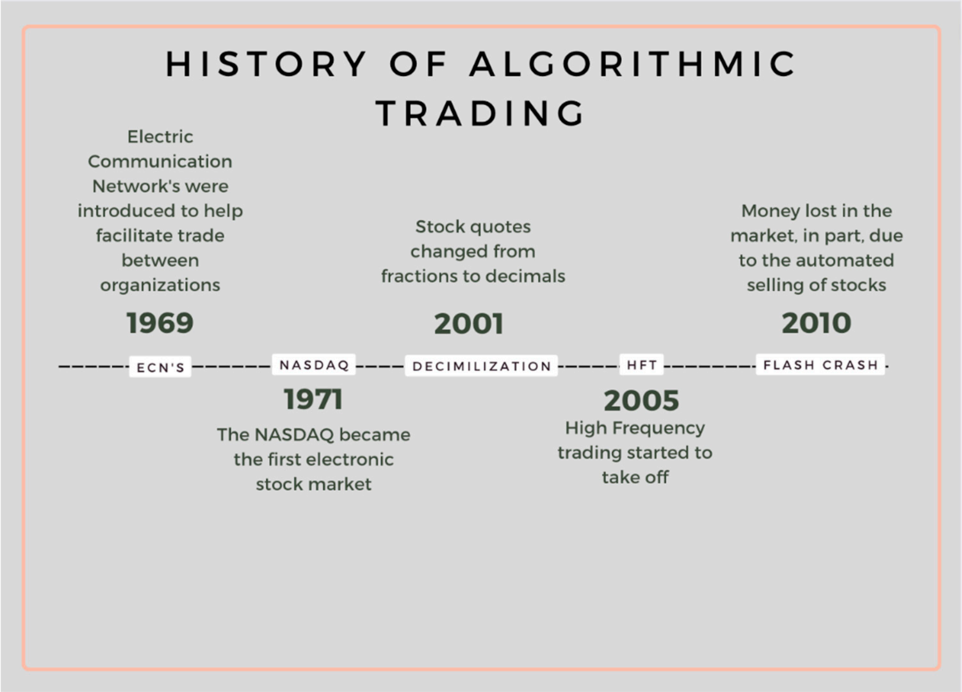
HISTORY OF ALGORITHMIC TRADING
Electric Communication Network's were introduced to help facilitate trade between organizations
1969
ECN'S
NASDAQ
1971
The NASDAQ became the first electronic stock market
Stock quotes changed from fractions to decimals
2001
DECIMILIZATION
HFT
2005
High Frequency trading started to take off
Money lost in the market, in part, due to the automated selling of stocks
2010
FLASH CRASH

Figure 2. History of algorithmic trading.

In 2001, decimalization, or the changing of stock quotes from fractions to decimals, created a more liquid market which, in return, helped with algorithmic trading, because the differences between buy and sell prices could be so much smaller. Then, the rise in algo trading was just beginning.

There are a few types of trading algorithms that have been developed. For example, there are algorithms and high-frequency trading. Hilbert and Darmon (2020) [20] provided the differences between these trading algorithms. They stated that algorithms simply automate trades that must be carried out. Algorithms do not necessarily create more profit, but they reduce risk. Duhigg (2009) [21] states that algorithmic trading, "consist[s] of instructions that allow humans to set the parameters for trade execution, such as a determined timeframe, volume patterns, risk-adjusted real-time market conditions, relative prices between selected stocks, etc.".

High-frequency trading (HFT) is a type of algorithmic trading that uses supercomputers to execute trades at very high speeds, and it is known for its very short holding periods. They create competition in the market, because there is an increase in the trade volume and a decrease in the time taken by the trades to be completed. HFT became popular in 2005, and between 2005 and 2009, it started to take off. According to the NYSE, HFT rose by 164% in that time period [21].

Algorithms and HFT are what some suspect led to the Flash Crash in 2010. The Flash Crash occurred on 6 May 2010, and the stock market lost $1 trillion, but then gained almost all of it back after 36 min. Many reasons have been speculated as to why this crash occurred, but the SEC stated that a company, Waddell & Reed Financial Incorporation, majorly impacted the crash. Waddell & Reed Financial Inc. sold $4.1 billion S&P 500 futures contracts using a sell algorithm. This meant that the contracts were sold in seconds without worrying about the time or price. HFTs that buy and immediately sell contracts force market makers to exit the market, and arbitrage traders buy cheap future versions of their stocks, leading to algorithms being completely confused and then to the Flash Crash of 2010. It is important to note that algorithms and HFT are not the only reasons for this crash, but is a good reminder of how algorithms can become confused with the data they are receiving. It shows that the market may not fully understand how algo trading systems work, and what triggers flash crashes.

3.2. How Data Science Is Used in Algorithmic Trading

For algorithmic trading to be beneficial, computers need to comprehend data sets that represent events in our fast-paced and changing world. However, help from human is still needed. First, humans develop a set of rules. These rules give instructions to the software on when to trade based on certain criteria. Some of these criteria are based on the price, volume, world events, historical pricing, and patterns. All of this information is put into an algorithm to determine what trades should be made.

This cannot be performed without the analysis of data sets. Data science helps with algorithmic trading by enhancing some of its common functions. It helps to collect, analyze, clean, and process data as well as identifying patterns and risk. While analyzing past and present patterns to make future predictions, data science allows for time series models, such as autoregressive models and exponential smoothing, to be completed. Data science also helps with back-testing, which is crucial for testing algorithms.

3.2.1. Time Series Models in Algorithmic Trading

Autoregressive models have past values that can predict what future values will be. According to Pennsylvania State University Statistics, "An autoregressive model is when a value from a time series is regressed on previous values from that same time series", for example, y_t on y_{t-1}.

$$y_t = \beta_o + \beta_1 y_{t-1} + \mathcal{E}_t \quad (1)$$

In this regression model, the response variable in the previous time period has become the predictor, and the errors have our usual assumptions about errors in a simple linear regression model. The order of an autoregression is the number of immediately preceding values in the series that are used to predict the value at the present time. So, the preceding model is a first-order autoregression [22].

The issue with using autoregressive models for algorithmic trading is that the model is very linear and cannot capture the complexity of financial market relationships. However, they have been very beneficial for predicting short-term trends in the market.

In addition to historical data, real-time data also play a crucial role in algorithmic trading. Real-time data can help algorithms to track market changes and immediate events and make trading decisions based on up-to-date information. Data science can absorb this information by building real-time data stream analysis models and combining it with historical data to provide more accurate trading decisions.

Exponential smoothing is another time series model that puts more emphasis on current data points than previous data. This is because it is believed that the more current data will be more relevant. Again, this method does not perform well with complex data, but it is still a good method for time series modeling.

Information that algorithmic traders use can be about the stock price, volume, current news, and earnings. This information is time-varying and can provide insight into market trends. By having time series models like autoregression and exponential smoothing, data on trends and seasonality can be found and analyzed.

In addition to traditional time series models, data science can improve the accuracy of algorithmic trading by applying more powerful machine learning algorithms. Deep learning algorithms can model nonlinear relationships to better capture the complexity of financial markets. Additionally, reinforcement learning algorithms can be used to train algorithmic trading strategies, so that they can self-adjust and improve in response to market feedback.

3.2.2. Back-Testing in Algorithmic Trading

Back-testing, which involves testing a strategy or algorithm using historical data, is how algo traders can see how their strategies would have performed in the past. Back-testing can provide statistics on volatility, net gain or loss, and annual return. Back-testing is needed in algorithmic trading to test how accurately a strategy performs. Without it, strategies might not be the most effective and may have theoretical flaws, which is not good, if they are used in real-world trading.

Data science helps to make back-testing successful by collecting historical data, testing the performance of the trading strategy, finding ways to make improvements to the strategies, and evaluating successes and issues with the back-tested trading strategy. Data science can automate the back-testing process to save time and provide high-quality analysis on market trading strategies. Overall, back-testing's, "...primary benefit lies in understanding the vulnerabilities of a strategy through a simulated encounter with real-world conditions of the past. This enables the designer of a strategy to "learn from history"" [18]. Algorithmic trading is focused on finding new models and strategies for improvements, and back-testing allows for the discovery and analysis of new models.

Algorithmic trading relies on data science to analyze past and present market data. Data science can analyze patterns in markets, clean data from outliers, create time series models to determine when to buy or sell, back-test to evaluate trading strategies, and make trades in an instant. Data science helps algorithmic trading continue to learn and improve over time and has helped to create a new form of stock market trading.

Data science provides tools and methods to assist in the decision-making process of algorithmic trading. This is important for regulation, compliance, and risk assessment. For example, the transparency and intelligibility of algorithmic trading can be enhanced by visualizing and interpreting the decision rules of machine learning models. Moreover, in algorithmic trading, data science can improve the quality of data through preprocessing techniques, such as removing noise, filling in missing values, handling outlier data, and selecting meaningful features.

In this process, it is worth mentioning that buying raw, especially cleaned data is very expensive, and cleaning the data is time-consuming, but it is essential, due to the sensitivity of the trading algorithm.

3.3. Trends of Algorithmic Trading

Algorithmic trading is expected to continue to grow. The International Market Analysis Research and Consulting Group's Algorithmic Trading Report of 2022 stated that, "The global algorithmic trading market reached a value of US$14.0 billion in 2022. Looking forward, IMARC Group expects the market to reach US$26.8 billion by 2028…" [23]. Algorithmic trading is already improving with the involvement of machine learning, which helps us learn from data and correct mistakes that were made. It will continue to help make HFT faster and more efficient, create better natural language processing skills to analyze world events that may affect the market, and remove human error in traditional trading. Although algorithmic trading can remove human emotion and error from the equation, humans are currently still needed to aid in market trading. "Results show pricing errors for NYSE stocks increase by 2–6% after floor trading is removed, illustrating how humans continue to be valuable even in the age of algorithmic trading" [24]. Humans still need to develop the strategies for algorithmic trading, they can react to non-forecasted events that an algorithm cannot, and they can provide moral and ethical considerations. For now, algorithmic trading will continue to grow, but humans will still play a large part in stock market trading.

Determining how to use generative AI in algorithmic trading is a potential research topic. Generative AI can be used to analyze market data and identify patterns. It can help develop more sophisticated trading algorithms. Reinforcement learning can be trained to make trading decisions based on historical data and real-time market conditions. However, it is crucial to address some challenges posed by generative AI, such as potential biases in data.

3.4. Opportunities and Challenges Associated with Algorithmic Trading

Algorithmic trading presents both positive and negative impacts. The future of algo trading can provide benefits by increasing the trading efficiency, introducing risk management techniques, and reducing the amount of work for traders. On the other hand, the challenges are still prevalent with market volatility issues, a lack of transparency, compliance with multiple regulations, ethical issues, and the possibility of data breaches and system failures.

3.4.1. Opportunities with Algorithmic Trading

Increasing Trading Efficiency

As mentioned before, HFT and algorithms will continue to increase the trading efficiency. The time taken to complete a trade will be significantly shorter, and these times will continue to improve. With more efficient algorithms, there will be reduced trading errors and lower transaction costs. The current world is all about speed and accuracy, and algorithmic trading helps market trading to comply with the current world standards.

Risk Management Techniques

Algorithmic trading allows for risk management techniques, such as stop-loss orders, take-profit orders, and portfolio diversification. Stop-loss orders automatically sell a stock, once it hits a certain price to prevent greater losses. They take the emotion out of trading and help an investor stick to a predetermined strategy. Take-profit orders are the opposite of stop-loss orders. They sell stocks when they hit a price that is higher than the original purchase price. Lastly, algorithmic trading allows for portfolio diversification, which aids in risk management. Portfolio diversification helps one invest in multiple things and have them all monitored. It takes the risk out of a wide scheme of investments. Your eggs are in multiple baskets.

Reduce the Amount of Work for Traders

Trading can create repetitive tasks for workers. Algorithmic trading can automate many repetitive tasks, such as watching screens, making trades, and monitoring markets in multiple time zones. This allows human traders to have more time to meet with clients, conduct research, and fulfill other tasks that humans must complete.

Reduce Workload for Traders

Trading may involve repetitive tasks, such as watching screens. Algorithmic trading can automate them.

Machine Learning for Market Movement Prediction

In algorithmic trading, it is generally enough to predict the direction of movement for a financial instrument. As such, forecasting whether there is an upwards or downwards trend becomes a classification problem. Machine learning, which is a branch of AI that involves the development of models that enable computers to learn from data and make predictions, could be used to predict the stock price [25]. Also, deep learning, which is the subset of machine learning that is based on artificial neural networks, could be used to forecast the market direction.

3.4.2. Challenges with Algorithmic Trading

Market Volatility

Market volatility is the amount of fluctuation that the stock market endures. This can lead to parameter sensitivity, because algorithms depend on set parameters to buy or sell. If the market is highly volatile, then these parameters of when to buy and sell need to be adjusted. Algorithmic high-frequency trading also creates market volatility because of the high number of orders being performed in seconds and because the stocks are only being held for seconds. The speed at which HFTs operate creates price fluctuations and unpredictability in the market. Market volatility is not all bad, but it creates challenges that algorithms need to stay on top of.

Lack of Transparency

Algorithmic trading relies on complex algorithms and fast execution. The nature of algorithmic trading can be confusing to the public, especially when algorithms are not public information. Market users are trusting their money with a mathematical algorithm. There are not always simple ways to understand why an algorithm is performing the way it is and the decision-making behind the trade. Humans can trust human-made decisions and, therefore, might choose a human trader rather than algorithmic trading. It is also hard for transparency and fairness to be achieved, because algorithmic trading relies on the fastest and most up-to-date technology. There can be price differences due to data transmission rates that can cause unfair trading.

Regulation Compliance Difficulties

The United States Security and Exchange Commission is developing regulations for algorithmic trading to be safe and positively affect the market. FINRA, the Financial Industry Regulatory Authority, has strict rules for algorithmic traders to follow. There are regulation practices that cover general risk assessments and responses, software/code development and implementation, software testing and system validation, trading systems, and compliance [26].

Ethical Issues

These include data breaches and system failures in the computer or software, programming issues, breaking into the system, and changes in algorithms.

3.5. How Robo-Advising Is Different from Algorithmic Trading

Robo-advising provides diversification advice to individual investors and does not require the intervention of human advisers [27]. It is a similar technology to algorithmic trading, because they both use algorithms to make decisions. However, there are key differences between the two technologies. Robo-advising is more tailored to individuals instead of huge trading firms. This more personalized approach allows individuals to state their goals for their finances and state how much risk they are willing to take. Robo-advisors will then monitor the market and buy or sell in order to maintain one's goals. They will create a portfolio, put one in stocks for a longer time, and rebalance at the end of the year to adjust the portfolio to certain weights. The computer model will determine asset allocations and decide the funds or investments for a given person.

Algorithmic trading and robo-advisors are two relevant FinTech tools that use algorithmic techniques in the investment decision-making process, but there are some differences. Algorithmic trading is more suitable for large trading firms and professional traders, who use complex algorithms to automatically execute trades and pursue fast and high-frequency trading. Robo-advisors, on the other hand, are more suitable for individual investors, which can provide personalized investment advice based on their financial goals and risk tolerance. Moreover, algorithmic trading relies on fast and advanced trading techniques that require low-latency execution and high-speed data transfer. Robo-advisors pay more attention to user friendliness and ease of use, usually provide services on web or mobile applications, and do not require users to have professional trading skills. Algorithmic trading faces some challenges in terms of transparency and fairness, as it relies on state-of-the-art technology and high-speed data transmission. This can lead to price differences and unfair deals. Financial regulators have certain regulations on algorithmic trading aimed at ensuring the safety and fairness of the market.

Overall, algorithmic trading and robo-advisors are financial technology tools that can be used for investment, and the choice of tool depends on an individual's investment objectives, risk tolerance, and technical skills [18].

3.6. Real-World Case Study

A real-world case study for algorithmic trading is the D. E. Shaw group. It is a multinational investment management firm founded by David E. Shaw in 1988 and based in New York City. The company is known for the development of complicated mathematical models and sophisticated computer programs to exploit anomalies in the financial market. It is a pioneer in computational finance and quantitative investing. The company's systematic strategies run on quantitative and computational techniques. It aims to identify statistically robust market inefficiencies based on practical knowledge of markets and advanced computational methods. As of 1 September 2023, the D. E. Shaw group had more than $60 billion in investments and committed capital [28].

4. Data Science for Fraud Detection

Fraud, which is defined by the Oxford dictionary as "Criminal deception; the using of false representations to obtain an unjust advantage or to injure the rights or interests of another" is becoming increasingly common in all areas of the world. Fraud in finance can occur through phishing, ransomware, identity theft, stolen cards, and insider attacks. According to the Global Fraud Report for 2022, phishing/pharming/whaling, card testing, and identity theft are the most commonly experienced fraud attacks [29]. Phishing is a way that attackers manipulate people to share private information with them. They perform this by sending legitimate looking emails, messages, and websites and obtaining passwords and banking information through them. Whaling is a type of phishing attack, but it is targeted at high profile people, such as CEOs. Pharming is when attackers use social engineering to redirect users to malicious websites, even when they accurately type in the web address. Card testing is when a criminal tries to determine whether a stolen card is still activated and will make small transactions on the card. Lastly, identity theft is when a person's personal information is stolen and is being used by a criminal for fraudulent activities. The Federal Trade Commissions stated that customers lost $8.8 billion (about $27 per person in the US) in fraud scams in 2022, which is more than double the amount in 2021 [30]. Fraud detection is vital to protect customers and companies. Fraud detection will help to detect any suspicious activity that may be fraudulent early on, hopefully before damage is done. Other than early detection, fraud detection allows for financial protection, builds customer trust, provides cybersecurity, and prevents insider threats.

4.1. How Is Fraud Detection Performed?

Fraud detection can be performed with biometrics, authentication and authorization, and with data science. Biometrics help with fraud detection, because they make it much harder to access an account. Biometrics use fingerprints, voice, and facial recognition to authenticate a user. These personal traits can be hard to replicate and, therefore, make it harder for hackers to access accounts. User authentication and authorization allow for improved fraud detection, because they again lower the chance that an attacker can access an account. Multi-factor authentication (MFA) requires a user to provide two or more credentials to prove their identity. It is much more difficult to acquire two credentials and, therefore, authentication and authorization ensure that an authorized user is who they say they are.

Data science is a massive help for fraud detection, especially for financial institutions. Data science provides real-time monitoring, predictive modeling, and pattern recognition. These techniques allow large volumes of data to be analyzed and studied to prevent fraud.

Real-Time Monitoring

Real-time monitoring is the continuous and constant monitoring of data. In this situation, real-time monitoring involves constantly looking at transactions and log-in locations. By constantly monitoring transactions, it is easier to identify anomalies in the data. Real-time monitoring allows for a close watch 24/7, without having humans work through the night.

Predictive Modeling

Predictive modeling is a way to predict risk through certain activities and transactions. These predictions can create a risk score and then help to identity fraud. With past and present data, a machine can use data science to predict future issues and then allows a company to create a plan. If a predictive model shows that fraud will increase through phishing attempts, a company can then decide to invest in phishing education and detection devices. Predictive modeling is a way to be proactive and try to prevent fraud before it happens.

Pattern Recognition

Data sets provide information on past and present transactions, and these data are analyzed constantly. These data sets allow for a vast number of patterns to be found, and data science can detect patterns that humans cannot. These patterns can show legitimate transactions but also detect fraudulent transactions. Pattern recognition allows for anomaly detection in data sets.

Anomaly Detection

Anomaly detection is a data science technique that flags transactions or activities that largely deviate from a user's history. In a financial transaction scenario, there is a user named Josh, who lives in New York City and normally uses his card to purchase groceries, eat at restaurants, and buy video games and shoes. Suddenly, his card is making transactions in Germany at make-up stores. This is a big deviation, or anomaly, in normal activities from Josh's normal transactions and therefore would lead to a fraud alert.

Determining how to use generative AI in fraud detection is another potential research topic. Generative AI can be employed to detect anomalies and patterns that are indicative of fraudulent activities in financial transactions. It can learn the normal behaviors of users and flag any deviations from the established patterns. However, it is important to address some challenges posed by generative AI, such as protecting sensitive financial data.

4.2. Opportunities and Challenges with Fraud Detection

Fraud detection is slowing the amount of cybercrime, but it is not able to identify it all. Fraud detection provides a wide array of positive opportunities, such as minimizing financial risks and allowing for more personal security. The positive outcomes are improving financial institutions and lowering stress levels in case something goes wrong. However, there are still challenges that are arising with fraud detection, such as keeping up to date with the changing cybercriminal techniques, false positives, and complying with all regulations.

Changing Cybercriminal Techniques

Cybercriminals are purposely constantly changing their attack methods. Hacking is a business to these criminals, and they will continue to develop new methods to work against any fraud detection measures. What started with spyware has led to phishing, man-in-the-middle, ransomware, and spear phishing attacks. Companies might develop a way to minimize risks from phishing attacks, but then they are being hit in the back with a brandnew attack. It is extremely difficult to predict what type of attack will occur next, and it is impossible to create fraud detection methods to prevent it all. Fraud detection algorithms need to be constantly changed to keep up with the changing cybercrime techniques. It is difficult and time-consuming to create algorithms that work properly. Every time a new type of attack is invented, an algorithm needs to be created to defend it. The cybercrime world is constantly changing and evolving, which causes fraud detection to be one step behind and it is responsive to what criminals do.

False Positives

False positives occur when fraud detection flags an event or action as fraudulent when in reality the action was legitimate. False positives create a hassle for the customer and company, because it creates more hoops to jump through. If a charge you make on your card is marked as fraudulent, your transaction and card can be canceled. This then creates a process of calling the bank and waiting for things to be sorted out. People may feel at ease that the bank is watching for anomalies and shuts down a card if they believe there are fraudulent activities; however, some people may find it troublesome that they have to go through hoops to obtain another card.

Compliance Regulations

Fraud detection systems rely on data to be able to detect anomalies and trends. Data support fraud detection systems, but storing personal data can be tricky. There are many regulations around storing personal data, such as the General Data Protection Regulation (GDPR), which is the EU privacy law and Section 5 of the United States Federal Trade Commission Act that supports personal data from companies. However, the United States is yet to create as in-depth regulations, like those that exist in other countries, but some individual states have taken that step. No matter where and what regulations are in place, it can be difficult to obtain personal data that are needed for fraud detection to work, due to the regulations. Fraud detection needs to collect data on personal information, such as names, addresses, financial statements, and physical locations to best detect fraud. This information can detect whether a purchase is completely out of the ordinary from their purchase history. However, these regulations on data privacy make it difficult for some systems to have access to that information. Not all systems can have access or the right to store personal information. Fraud detection systems may have the best security in mind, however, there are many regulations to be followed to protect the people.

Ethical Challenges in Financial Data

There are several challenges associated with the handling of financial data. First, financial data often contain sensitive information about individuals' incomes, spending habits, and financial histories. It is crucial to maintain the privacy of these data. Second, financial institutions need to ensure data security. If data breaches happen, it may cause financial losses for individuals and reputational damage for institutions. Third, misleading or inaccurate data can lead to poor financial decisions for investors. It is essential to ensure the accuracy and integrity of financial data.

4.3. Financial Data Sets

Financial data sets are crucial to allow analysts and traders to make informed decisions. Indeed, it is the quality of the data sets that matters, not whether there are big or small data. As an example, Roccetti et al. (2020) [31] trained a neural network with over fifteen million water meter readings, but it failed to predict when a meter would malfunction. So, the authors only used the samples that were free of noise for training, and the neural network eventually achieved a prediction accuracy of over 80%. This illustrates that it is important to clean and reorganize noised datasets, so that they can describe the underlying statistical phenomenon.

As for financial fraud detection, relevant data sets for analysis are credit ratings for individuals and companies; corporate financial statements that include income statements, balance sheets, and cash flow statements; and other useful data. As for algorithmic trading, relevant data sets for research are stock market data that include historical stock prices, trading volumes, earnings and dividend data; economic data that include Gross Domestic Product (GDP) data, unemployment rates, the Consumer Price Index (CPI), and interest rates; and social/news data that include sentiment data from social media, RSS feeds, and news services [18].

5. Conclusions

Data science is becoming increasingly popular in multiple fields, due to its ability to extract meaningful insights from data. Particularly in the finance sector, data science has allowed the financial technology industry to soar. Data science is allowing financial technology companies to collect, analyze, clean, and draw insights from data, which allows them to better serve customers. Data science in finance has allowed for algorithmic trading, robo-advising, fraud detection, etc. Also, it has applications in credit scoring, loan approval, and mortgage rate forecasting. We plan to explore these topics in the future.

The opportunities that data science has created in these areas have been immense. Data science has allowed us to use digital and international payments. It has allowed for algorithmic trading to be more efficient and have lower risks. Overall, it provides a more personal, safe, and efficient way for financial tasks to be completed. However, the challenges can be difficult to manage. The ways that data science has helped with FinTech, algo trading, and fraud detection have also created issues with regulatory compliance, personal security, ethics, and financial inclusion. There are plenty of benefits to data science in finance; however, there are still challenges that hinder people from fully supporting or using the resources that data science in finance provides. Information and data are gold, and it is important to mine the gold, using data science to understand its value.

Author Contributions: Conceptualization, X.Z.; writing—original draft preparation, E.G.; writing—review and editing, X.Z.; supervision, X.Z.; resources, S.C., T.Z., S.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Easwaran, B.; Aramuthakannan, S.; Lokesh, S.; Kumar, R.; Saikia, K.J.; Saikia, U. A study on the use of data science in healthcare applications and the mathematical issues in data science. *J. Algebr. Stat.* **2022**, *13*, 2535–2541.
2. Chakravaram, V.; Vidya Sagar Rao, G.; Jangirala, S.; Sunitha, R. The role of big data, data science and data analytics in financial engineering. In Proceedings of the International Conference on Big Data Engineering, Hong Kong, China, 11–13 June 2019; pp. 44–50.
3. Özsu, M.T. Data science—A systematic treatment. *Commun. ACM* **2023**, *66*, 106–116. [CrossRef]
4. Cao, L. Data science: Challenges and directions. *Commun. ACM* **2017**, *60*, 59–68. [CrossRef]
5. Cao, L. AI in finance: Challenges, techniques, and opportunities. *ACM Comput. Surv.* **2022**, *55*, 1–38.
6. TensorFlow. TensorFlow Case Studies. 2023. Available online: https://www.tensorflow.org/about/case-studies (accessed on 20 November 2023).
7. NLTK. Natural Language Toolkit. 2023. Available online: https://www.nltk.org/ (accessed on 20 November 2023).
8. IBM. What Is Data Science? 2023. Available online: https://www.ibm.com/topics/data-science (accessed on 20 November 2023).
9. Apache. The Apache Hadoop Software Library. 2023. Available online: https://hadoop.apache.org/ (accessed on 20 November 2023).
10. Barbaglia, L.; Consoli, S.; Manzan, S.; Recupero, D.R.; Saisana, M.; Pezzoli, L.T. Data science technologies in economics and finance: A gentle walk-in. In *Data Science for Economics and Finance: Methodologies and Applications*; Consoli, S., Recupero, D.R., Saisana, M., Eds.; Springer: Cham, Switzerland, 2021.
11. Pisoni, G.; Molnár, B.; Tarcsi, Á. Data science for finance: Best-suited methods and enterprise architectures. *Appl. Syst. Innov.* **2021**, *4*, 69. [CrossRef]

12. TransUnion. Fraudsters Shift Focus at Mid-Point of 2021 from Financial Services to Travel and Leisure and Other Industries. 2021. Available online: https://newsroom.transunion.com/fraudsters-shift-focus-at-mid-point-of-2021-from-financial-services--to-travel-and-leisure-and-other-industries/ (accessed on 20 November 2023).
13. Das, S.R. The future of fintech. *Financ. Manag.* **2019**, *48*, 981–1007. [CrossRef]
14. Hendershott, T.; Zhang, X.; Zhao, J.L.; Zheng, Z. Fintech as a game changer: Overview of research frontiers. *Inf. Syst. Res.* **2021**, *32*, 1–17. [CrossRef]
15. Grand View Research. Artificial Intelligence in Fintech Market Size Report. 2023. Available online: https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-in-fintech-market-report (accessed on 20 November 2023).
16. Tok, Y.W.; Heng, D. Fintech: Financial Inclusion or Exclusion. 2022. Available online: https://www.imf.org/en/Publications/WP/Issues/2022/05/06/Fintech-Financial-Inclusion-or-Exclusion-517619 (accessed on 20 November 2023).
17. Kharpal, A. China's ICBC, the World's Biggest Bank, Hit by Cyberattack that Reportedly Disrupted Treasury Markets. The New York Times. 2023. Available online: https://www.cnbc.com/2023/11/10/icbc-the-worlds-biggest-bank-hit-by-ransomware-cyberattack.html (accessed on 20 November 2023).
18. Treleaven, P.; Galas, M.; Lalchand, V. Algorithmic trading review. *Commun. ACM* **2013**, *56*, 76–85. [CrossRef]
19. Mukerji, P.; Chung, C.; Walsh, T.; Xiong, B. The impact of algorithmic trading in a simulated asset market. *J. Risk Financ. Manag.* **2019**, *12*, 68. [CrossRef]
20. Hilbert, M.; Darmon, D. How complexity and uncertainty grew with algorithmic trading. *Entropy* **2020**, *22*, 499. [CrossRef] [PubMed]
21. Duhigg, C. Traders Profit with Computers Set at High Speed. The New York Times. 2009. Available online: https://www.nytimes.com/2009/07/24/business/24trading.html (accessed on 20 November 2023).
22. Pardoe, I. Autoregressive Models. 2023. Available online: https://online.stat.psu.edu/stat501/book/export/html/996 (accessed on 20 November 2023).
23. IMARC. Algorithmic Trading Market Research Report. 2023. Available online: https://www.imarcgroup.com/algorithmic-trading-market (accessed on 20 November 2023).
24. Manne, K. Not So Fast, Robots: Humans Are Still the Best Stock Traders. 2023. Available online: https://www.buffalo.edu/news/news-releases.host.html/content/shared/mgt/news/not-so-fast-robots-humans-still-best-stock-traders.detail.html (accessed on 20 November 2023).
25. Koshiyama, A.; Firoozye, N.; Treleaven, P. Algorithms in future capital markets: A survey on AI, ML and associated algorithms in capital markets. In Proceedings of the First ACM International Conference on AI in Finance (ICAI), New York, NY, USA, 15–16 October 2020; pp. 1–8.
26. FINRA. Algorithmic Trading. 2023. Available online: https://www.finra.org/rules-guidance/key-topics/algorithmic-trading (accessed on 20 November 2023).
27. D'Acunto, F.; Prabhala, N.; Rossi, A.G. The promises and pitfalls of robo-advising. *Rev. Financ. Stud.* **2019**, *32*, 1983–2020. [CrossRef]
28. Shawn, D.E. Who We Are. 2023. Available online: https://www.deshaw.com/who-we-are (accessed on 20 November 2023).
29. Cybersource. Global Fraud and Payments Report 2022. 2022. Available online: https://www.cybersource.com/content/dam/documents/campaign/fraud-report/global-fraud-report-2022.pdf (accessed on 20 November 2023).
30. Federal Trade Commission. New FTC Data Show Consumers Reported Losing Nearly $8.8 Billion to Scams in 2022. 2023. Available online: https://www.ftc.gov/news-events/news/press-releases/2023/02/new-ftc-data-show-consumers-reported-losing-nearly-88-billion-scams-2022 (accessed on 20 November 2023).
31. Roccetti, M.; Delnevo, G.; Casini, L.; Salomoni, P. A cautionary tale for machine learning design: Why we still need human-assisted big data analysis. *Mob. Netw. Appl.* **2020**, *25*, 1075–1083. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.