

# Chapter 1

## Artificial Intelligence and Machine Learning in Financial Services to Improve the Business System



Komalpreet Kaur, Yogesh Kumar, and Sukhpreet Kaur

**Abstract** Machine learning is coming as a significant encroachment in the financial services industry. Finance has always been about data and is considered a complex field of study that includes knowledge from disciplines such as mathematics and statistics to human psychology and linguistics. Due to this, it is difficult to manage the various day-to-day challenges associated with finance, such as financial glitches attributed to human errors. The financial sector has employed machine learning for a myriad of purposes and its excellent applications. The work highlights the advancement of different learning techniques in financial services for data science. This chapter gives comprehensive prospects on a study accomplished in the financial industry after applying digital financial solutions over time and an extensive view of this distinctive research area. The organization comprises an introduction, motivation, and background that entails a block diagram of learning techniques, benefits, and various issues related to implementing machine learning techniques in the financial domain. Other things covered in this chapter are various datasets used by different researchers, and its focal point is to present a systematic survey of various applications of finance using artificial intelligence and finally expose a synthesis analysis based on the findings along with their benefits and issues. Overall, this chapter gives conscious and constructive assistance to researchers working towards the sustainable evolution of the finance industry.

---

K. Kaur

Department of Electronics and Communication Engineering, Punjabi University, Patiala, India

Y. Kumar (✉) · S. Kaur

Department of Computer Science and Engineering, School of Technology, Pandit Deendayal Energy University, Gandhinagar, Gujarat, India

e-mail: [yogesh.kumar@sot.pdpu.ac.in](mailto:yogesh.kumar@sot.pdpu.ac.in); [Yogesh.arora10744@gmail.com](mailto:Yogesh.arora10744@gmail.com)

S. Kaur

e-mail: [er.sukhpreetkaur@gmail.com](mailto:er.sukhpreetkaur@gmail.com)

S. Kaur

Department of Computer Science & Engineering, Chandigarh Engineering College, Landran, Mohali, India

**Keywords** Data science · Financial services · Machine learning · Artificial intelligence · Deep learning · Datasets · Accuracy

## 1 Introduction

The data evaluation is based on test and error, which is an outlook that becomes unpractical in the case of large and diverse data sets. Machine learning (ML) comes as a besotted way to evaluate this massive amount of data. This rapid progress in ML increases its usage, demands, and importance in contemporary life. This has also replaced data mining in which explications are done using automating generic methods that have replaced the conventional statistical methods. So, precise results can be created using ML by evolving proficient and speedy algorithms and data-driven models for real-time processing [1]. The finance industry is broad and different segments have different use cases for ML. The use cases for ML in finance are both abundant and highly important. Fraud detection, credit quality assessment, algorithmic trading, and compliance regulation in real-time are examples of use cases [2]. ML in finance has refolded the financial services industry like never before. It helps businesses modify their customer experience and enables them to offer personalized products and services based on consumer actions [3]. ML techniques forecast and help diminish customer dissatisfaction, like an exact prediction of cash required in ATMs helps manage costs and improves the return on cash assets. Chabot offers 24\*7 customer service with superhuman speed and constancy. ML underwriting agents assist human underwriters to make the best utilization of their time and effort. Fraud detection algorithms help find suspicious activity that the best human experts may fail to spot. Even business association is not wholly human activity any longer. Machine learning makes specific that these digital financial solutions persist in to carry out correctly, even as the wants of financial institutions grow significantly over time [1]. Due to its popularity in various fields, various researchers have also worked on machine learning in the finance sector. However, still, they were unable to give reviews on things related to it. Many researchers have presented the use of machine learning or artificial intelligence in the finance domain and presented using the technique for individual finance. However, there is no good survey or research presented in the area of artificial intelligence covering its different aspects and its applications in the field of finance, the role of data science in finance, and comparative analysis of previous research done in this field. There is a need for comprehensive study in the field of machine learning especially in the domain of finance. So, the objective and motivation behind this chapter are to outline the use of ML in the finance domain and the challenges faced while implementing it, and various advantages achieved by applying both together. Along with it, highlights a systematic review of various datasets or financial data used by machine learning algorithms to carry out the experiments on it.

## 1.1 Organization of the Chapter

The complete chapter is organized into various sections for terms related to the Finance sector. The sections are divided into numerous sections according to the main terms related to the use of AI in Finance. Started with Sect. 1 which contains the introduction of finance and the use of AI in it forwarded with the motivation behind writing this chapter given in Sect. 2 which also consists of challenges faced by the user while using AI in the finance sector. Further, in Sect. 3 Background is given about the main topic that contains data related to the role of data science in the finance sector, benefits and issues of AI in the finance sector, datasets used in financial applications along how AI is changing the financial services industry (Fig. 1).

Then, Sect. 4 contains the role of AI in the finance sector is included that further divided into various subheadings of applications of finance in which AI plays its role. The comparative analysis is also given in Sect. 5 in which a comparison of various applications of finance has been done in terms of the dataset used along with output achieved using different AI algorithms.

## 2 Motivation

Finance has always been about data. AI is used for digesting immense amount of data and learning from it by carrying out a particular task such as differentiating fraudulent legal documents from all authentic documents. Thus, ML learns from presented data to divide future and data-driven decision-making models from input data sets [1].

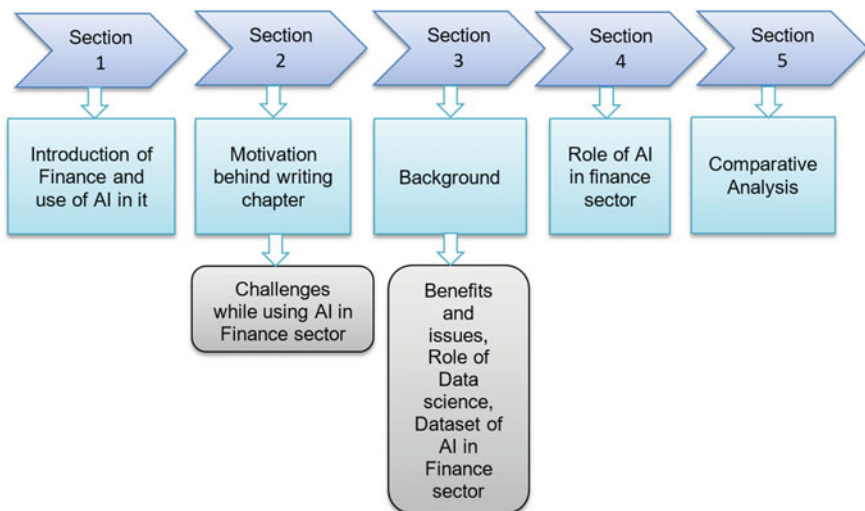


Fig. 1 Organization of the chapter

AI in finance gives numerous approaches for handling large and complex amounts of information in a better way. It also improves the ability of managing such large and complex data accessible by the finance industry [3]. Due to a large amount of historical financial data produced in the industry, machine learning has found many constructive applications like trade settlement, fraud detection, algorithmic trading, high-frequency trading, chatbots, loan underwriting, Robo-advisors, risk management, money laundering, and document analysis. AI using the latest technologies helps in different ways in the finance sector, from approving loans and executing credit scores to handling assets and estimating risk [4]. No doubt, machine learning has taken the finance industry to an all-new level. However, still, it still faces some challenges in its implementation in the industry like cost, security, updates, adopting challenges, etc. [2]. Along with these challenges, machine learning has various advantages like time-saving, less paperwork, improved workflow, etc., that attracted us to work on it. Moreover, its use in various areas to engross researchers for working on it improves the existing systems. It has been found that acceptable improvement is achieved in various areas of finance by using machine learning.

## 2.1 Challenges

As machine learning is enjoying a moment of resurgence, there are implementation challenges a finance sector should tackle to be successful. There are many challenges that are faced while implementing machine learning in the finance sector. Such challenges are shown in Fig. 2 [2, 5, 6].

**Data-**Banking or financial data is often poor quality and hard to find as it is stored in silos on various legacy systems. Algorithms flourish easily accessible large data sets. The assimilation of data sources, ideally onto a cloud platform, is hence a key.

**Regulatory-**Some self-learning models cannot be conventionally validated and consequently may be considered insufficient by the regulator. Comprehensive research into regulatory requirements is suggested in advance of implementation.

**Tools-**There exists an immense array of new and growing Machine Learning technologies. A systematic consultation course of action with digital specialists is advised at the forefront of any purchase.

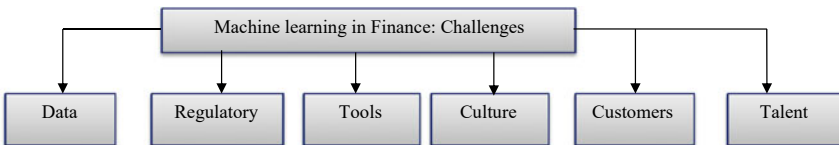


Fig. 2 Challenges of machine learning in finance

**Culture-Judgment** currently often trumps insights in firms- a cultural shift will be required. The democratization of analytics is required, and there should be a lure to sway data sharing between business divisions.

**Customers**-Older generations are not so much digital-savvy. Customers choose human interaction over communication with robots. A marketing /education session may be required to emphasize the benefits to the customer.

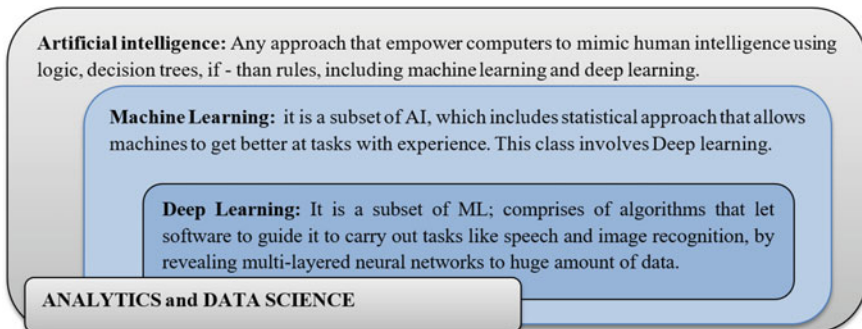
**Talent**-Initiating machine learning into business needs a change in skillset requirements from operational manager to analytics and data science.

### 3 Background Study

This section expounds on a brief overview of learning techniques and the role of data science in financial domain applications, along with its advantages, challenges, and objectives. Various application areas of machine learning in finance are also presented in this section.

Conventional computer science algorithms were application-specific, and time was necessary to develop a proficient algorithm for a specific application. Each application generally had different requirements, and as a consequence, there were not even single approaches that fit all methods. Artificial Intelligence (AI), ML, and Deep Learning assure a subset of algorithms that fit a range of sub-tasks under a specific application sphere [4].

Based on skills in using computers for surveying the data for structure, ML has evolved even in the cases of having no presumption about the structure. The experiments on it are a validation error on new data. On the other hand, a null hypothesis is proved by theoretical experiments. This happened due to frequent use of the iterative approach by ML for learning from data and automating the learning. Until a robust pattern is found, passes are run through the data [7] (Fig. 3).



**Fig. 3** Definition of various analytics and Data Science learning [4]

In the field of Deep Learning (DL), which is a subset of Machine Learning, the patterns of large amounts of data are studied. In DL, for dealing with data, neural networks are used which in turn also increases the computing power of the system. The current applications of DL include object recognition in images as well as the word recognition in various sounds. It is also being used in translating different languages automatically, medical diagnosis and other problems related to social issues by all the researchers [8]. That is why, the field of Data science is referred to as the area in which knowledge is extracted from huge data sets having data of unstructured and structured types. This knowledge is further being shared with their domains of business so that they can offer effective schemes and roadmaps in their business [9].

### 3.1 Role of Data Science in the Finance Sector

Industries observe data as a crucial commodity and energy. Raw data is used in it which is further converted into a meaningful product. This insight the enhancement in industry performance and another name of finance is the hub of data and its industries are considered as one of the preliminary users and explore data analytics. In fraud detection, customer management, algorithm trading and risk analytics data science is used [9, 10] (Fig. 4).

- Risk Analytics: It has an important place in different applications of finance related to data science. By using it, deliberate decisions can be taken by the company along with the rise in reliability as well as in-company security. In risk management, a core is data because it estimates the rate of loss recurrence and multiplies it with the magnitude of damage. There are different types of risks, which are evaluated by a company that arise from competitors, credits, market, etc. In managing risks, the primary step is its recognition along with

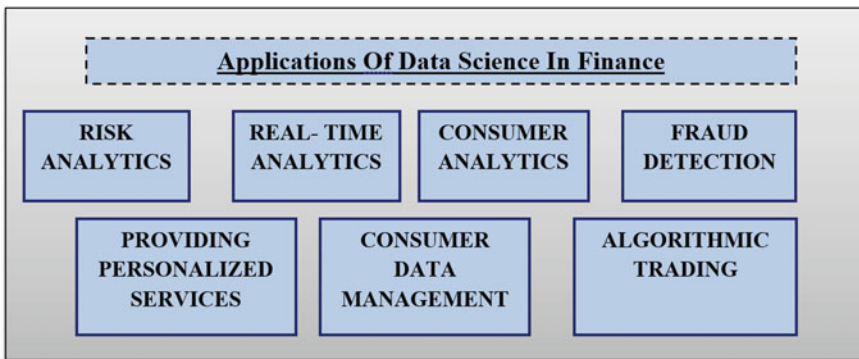


Fig. 4 Data Science for difference finance applications

its observation and prioritization. The data relating to customer information in the financial transaction is available in a massive manner. Due to it, institutions are trained on risk data, so that the risk scoring models can be implemented optimally with less expense. The creditworthiness of customer authentication is another important feature of risk management. For this purpose, data scientists are paid by companies that use machine learning algorithms to examine customer transactions [11, 12].

- **Real-Time Analytics:** Earlier the processing of data was performed in batches, which means it is chronological, not real-time. However, to get insights into present conditions, real-time data is required by most industries. That is why they faced problems. However, now with minimum latency, data can be accessed which happens with the evaluation of dynamic data pipelines and progress in technologies. In institutions, finances, this data science application can trace transactions; and generate credit scores along with other financial attributes without causing any latency problem [13].
- **Consumer Analytics:** Financial institution's primary function is consumer personalization. Data scientists are able to make suitable decisions and insights from consumer behaviour using the data in real-time analytics. The customer lifetime value is computed using consumer analytics by insurance companies like financial institutions. This will be added to their cross-sales that minimizes below zero customers for rationalizing the losses [14].
- **Customer Data:** Management financial Institutions require data. The working style of financial institutions has been transformed to much extent with the introduction of big data. Most of the data is provided through transactions and social media.

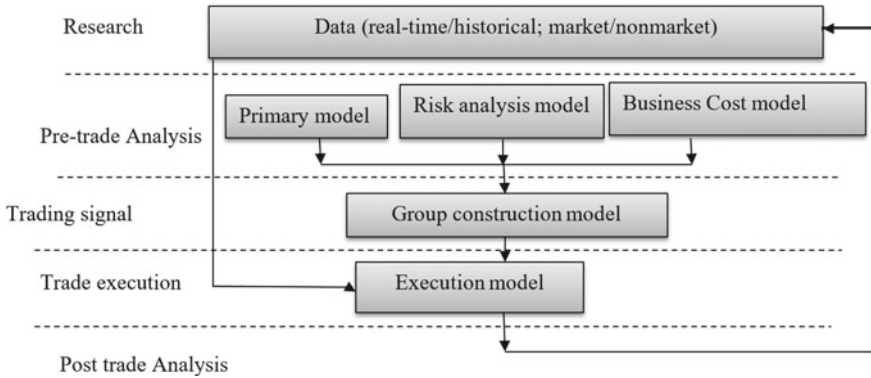
This data is presented in two forms as given below:

- Unstructured data
- Structured data.

Unstructured data causes many problems, whereas structured data is easier to handle.

The most significant part of Big Data is Business Intelligence (BI). BI helps industries in finding the important information about customers by applying machine learning. There are various other tools also which are used in AI for extracting meaningful information from the input data which includes data mining, text analytics, etc. By detailed analysis of customer data, the financial trends and market value changes are examined by machine learning [9].

- **Providing Personalized Service:** Personalized services to consumers are the liability of financial institutions that analyse consumer information using various techniques and further understand their interactions. Along with providing better communication to its users, financial institutions depend on software based on speech recognition and natural language processing. There is an increment in profits by taking actionable insights from financial institutions according to the



**Fig. 5** Components of algorithmic trading [18]

customers' feedback data. This also optimizes the strategies of industries, which in turn helps them in improving their services [9].

- **Fraud Detection:** The industry of finance is majorly affected by frauds, and with an increase in several transactions, there is a rise in fraud risks. However, now the financial industries can keep track of fraud by introducing and increasing big data and analytical tools. In these industries, the most common fraud is credit card fraud. By advancing in algorithms, this type of fraud comes to know as it has increased the reliability of fraud detection that alerts companies about particularities in the financial purchase, ultimately encouraging them to block the account and reduce the losses [15, 16]. The unusual patterns of trading data can be spotted by various machine learning tools and the aware financial institutions to explore it more. Some insurance-related frauds are dealt with by banks [17].
- **Algorithmic Trading:** The components of algorithmic trading are shown in below Fig. 5. As can be seen there are three stages of algorithmic trading predate analysis, Trading signal generation, and Trade execution. Predate analysis takes in the mathematical model [18].
- To trade financial instruments, future activities are visualized by the risk alpha model.
- The risk model estimates the financial instruments-related risk levels.
- The transaction cost model computes the financial instruments' trading-related costs.

The portfolio construction model is included in trading signal generation, and its input is the results of alpha, risk, and transaction cost models. It further decides the amount of financial instruments portfolio allowed to go onward and in how many quantities.

The trades are executed at trade execution that makes various decisions by checking the transaction expenses and trading time. The trading strategy accompanied by venue and order type is the most general decision.



In financial institutions, the most significant part is played by algorithm trading which includes complex mathematical formulas and high-speed computations. This helps financial institutions develop new trading strategies, and it has a massive impact on high-speed computations. By this, data science is considered an essential feature of financial institutions.

There is a massive collection of multiple data operations in data science that involve statistics and machine learning reliant on data. Then further in the form of training and test set, these data are fed to a model that helped in implementing the module with algorithm parameters. This states that the future of Data Science is dependent on the advancement in Machine learning, [9]. Data science also includes [10]: Data Engineering, Automating Machine learning, automated data-driven decisions, Data Integration, Data visualization, Dashboards and BI and Automated data-driven decisions.

### ***3.2 Benefits and Issues of Artificial Intelligence in Finance Sectors***

This section covers various benefits and issues of using artificial intelligence in finance sectors. The logic to implement artificial intelligence in business seems evident. Nothing but machines can offer the following benefits:

- Process automation has reduced operational costs [4].
- Better productivity and enhanced user experiences have led to increased revenues [19].
- Better compliance and Strong security [12, 20].
- Cost-effective.
- Increasing Operational efficiency.
- Improve security.

Companies that are providing financial services can use a large number of AI algorithms that are open source. Also, these companies are also ready to spend any amount of money for purchasing state-of-the-art hardware which is required for increasing the computing power of the system. Various areas of the industries related to finance are improved with the application of AI that helped in dealing with large amount of existing data [21].

ML and deep learning algorithms in finance reduce labor costs by automating human labor, resulting in a significant saving of money in financial services. Operational efficiency is also improved by streamlining processes that increase the productivity and efficiency of financial operations. It also provides both network fraud prevention and security capabilities for financial institutions. The bank also benefits from the model's ability to select the financial indicators that are most relevant in the process of prediction and a high level of prediction accuracy. Human capabilities

are improved, and AI's impact on business and the economy will be reflected in its direct contribution and ability to inspire complementary innovations.

The current development in AI mainly focuses on reducing the prediction cost and making the system faster and more accurate. These predicting action implications enable increased customer retention and prevent downtime through predictive maintenance on infrastructure or machinery. The benefits of AI are the ability of trades to execute at the best possible prices, increase accuracy, and reduce the mistakes likelihood with the ability to automatically and simultaneously check various market conditions. Human beings' emotional and psychological conditions cause some errors that also get reduced by replacing human work with AI systems. We have seen that various advantages are provided by AI technology, but still, some issues need to be addressed. Most of the financial services companies are not taking genuine value from AI technology for the reasons given below [2, 19].

- Research and Development in machine learning are expensive.
- The lack of machine learning engineers is another major issue.
- Selection of datasets for the experimentation.
- Financial incumbents are not smart enough when it comes to upgrading data infrastructure.

Sustainability and overfitting are two main issues in using AI models. Another possibility of an increase in the count of frauds using new approaches also increases the count of legal issues caused by mistakes of an algorithm. This also increases the count of privacy issues. Nowadays, AI models are used for preventing fraud, but there can be a possibility that the use of these tools will increase the number of cybercriminals defrauding users in the coming time. Also, the institute is legally responsible for whose person makes a mistake that also invokes legal questions on mistakes done by algorithms. The data is the primary source needed by these algorithms. However, the financial data is private information, so AI increases the concerns related to privacy.

### ***3.3 Datasets Used in Financial Applications***

In different studies or different applications of the financial domain, we use different datasets. A few of the datasets are available publicly that can be downloaded, such as Australian, Japanese, German, Korean, etc. These datasets are highly used in the prediction of bankruptcy and the scoring of credits. Some researchers have used publicly available data as well as some have collected data from specific countries also [22]. Table 1 shows some of the datasets and their description.

**Table 1** Description of various datasets [22–25]

Name of dataset	Description
Quandl	For economic and financial data their dataset is considered as a good source that is mainly proving to be beneficial in model construction for stock prices or forecast economic indicators
World Bank Open Data	Their datasets enclose the analysis of population as well as a large number of economic and development indicators worldwide
IMF Data	This dataset has all data associated with international finances, rate of investment, rate of debts, and reserves dealing in foreign exchange
Financial Times Market Data	Provide updated details on financial markets worldwide, together with stock price indicators, property, and foreign exchange
Google Trends	They did a study and investigation of internet search activity data along with the different stories trending across the world
American Economic Association (AEA)	This data set contains the macroeconomic data of the United States of America
EU Open Data Portal	It was started in 2012 following European Commission Decision 2011/833/EU on commission document reuse. This dataset provides the required data without any cost and it contains the data of European Union Institutions. The data present in this dataset is collected from varying datasets and can be used anywhere in any sector
EOD Stock prices	EOD stands for the end of day and it consists of around 3000 US companies end of day stock prices, splits, and dividends curated by the Quandl community
Data.gov	It is a website of the US government that gives access to high-value, machine-readable datasets of various domains generated by the federal government’s executive branch. The Vivek Kundra US federal CIO launched this website in 2009 that is powered by WordPress and CKAN open source applications
Global Financial Data (GDF)	The complete unabridged data series was generated by GDF that contains the combination of daily market data taken from historical and data feeds values. This can be accessed by anyone due to its free subscription and further can research on analysing the major global markets and economies
RBI Database	This dataset was launched by the Reserve Bank of India, RBI Data Warehouse which is a publisher’s data platform for various Indian economy aspects. In this, the data is presented through a time series formatted report. It contains a dataset across financial markets, saving, money, employment, banking, national income, and others
PBGC Financial Summary Data	This consists of PBGC’s key financial statements highlighted data of the period September 1992 to 2009. These statements can be accessed at PBGC.gov
Large Mutual Fund Facilities	This dataset contains a combined asset in US mutual funds that was \$12.36 trillion in Oct 2007. There are around 8,000 US mutual funds that are significantly concentrated assets in a relatively small amount of mutual fund families which is 50% of all assets held by Top 10 mutual fund families

(continued)

**Table 1** (continued)

Name of dataset	Description
US Stock Data	It contains US stock historical data since 2009 that is updated on a daily basis
CBOE Volatility Index (VIX)	It is a critical indicator of S&P's market forecasts for near-term volatility. This is a time series dataset that includes daily lows and highs, as well as closes and opens
Dow Jones Weekly Returns:	It contains a percentage of return that stock has every week for training their algorithm and measuring while producing the greatest rate of return by stock in the following week
Simin	The data provided by this dataset is available in one document so that it can be used in a timely manner. The data is collected from different financial statements. It is uploaded on the SEC website after its cleaning and organizing
CIA World Factbook	It includes the economic state of countries along with other military, communications, geography, and demographics stats

### ***3.4 How Artificial Intelligence (AI) is Changing the Financial Services Industry?***

In collecting and analysing data, AI is able to give much accurate future predictions and is very efficient in recognizing patterns in comparison to human ability. This helps in improving the effectiveness of banks in their routine work and helps in completing the tasks in less time. According to the remarks of the PWC's latest study, AI holds around 16 trillion dollars of the global economy in 2030. Moreover, its applications will cross 5 billion dollars of global investment, and by 2030 it is expected to save 1 trillion dollars in the banking industry [20]. There are various AI applications in the service industry. A few of them are given below.

**Dedicated Services to HNI's (High Net Worth Individuals) Wealth and Portfolio Management:** The main work of these services is to decide the trade-off between risk and return and in turn give warning to the users detailing various securities and assets with their possible returns. The advancement is given by AI to financial services companies which helps them give accurate and customized guidance to their well-off clients. BlackRock has helped the AI lab in its operations. It has been considered as the largest group of the world with higher investments having assets of 6 trillion dollars. Other global organizations are also using AI for improving predictions which is helping their clients.

Moreover, two AI systems have been brought by Swiss Bank UBS to renovate their trading floor. In this, after interpretation of market data, scope trading patterns can be found out by one. After that, it advises trading strategies to the bank's customers to help them achieve high returns. On the other hand, their customer's post-trade allotment preferences are conveyed by the second one [23].

- **Virtual Financial Assistance and Automated Customer Support through Robo Advisors and Chatbots:** The AI associate and other appropriate apps like Revolut are used by Banks that help in having quick services to the customers. For this purpose, they use Smart chat technologies that will help in transferring the queries of the customers to the supporting staff related to that query. It will require different processes of Natural Language Processing (NLP) for achieving this feature. After that, in 2016, an AI assistance Luvo was introduced by the Royal Bank of Scotland that answers customers' inquiries, and in some cases, it transfers them to human staff. This Robo advisor helps in reforming the experience and pleasure of customers. In India, the four top commercial banks are using Chatbots which are one of the major applications of AI. Along with it, they have advanced the customer experience by FinTech startups that reduce the cost and give better efficiency. On the other hand, to provide immediate feedback according to the captured facial expressions of clients, AI power-driven intelligent cameras are used by banks [24, 26, 27].
- **Enhanced Insurance Experience:** There are many applications for both instance of claim payment and underwriting policies in the data-driven insurance industry. The main requirements of insurance companies are to know more about the client's education, lifestyle, health, and character along with the filed claim occasion. However, these all things can be effectively captured using AI algorithms. In some US startups, there is a possibility of paying insurance claims within 3 s by doing multiple back-end procedures and checks using AI apps. They can check things while communicating with the customer at the front end [20, 28].
- **Robotic Process Automation (RPA)—Repetitive Task Automation:** There is a need for repetitive front and middle office processing in repetitive activities like deposit and withdrawal processing, billing, statement generating, cheque clearing, etc. However, with RPA and AI software, this can be accomplished better, resulting in better efficiency, improved time management, and expense savings. Human intelligence and skills are progressively imitating robotics technology from industrial robots to self-driving cars, which may become a game-changer in the financial services industry. There is speedy growth for investment in the robotic sector, and there is almost double the number of funding deals in robotics. According to CB Insights, in 2014 it was 273 million dollars, then in 2015, it became 587 million dollars. On the other hand, the expansion of investment was 115% in 2015 compared to 55% in 2014 [4, 20].
- **Use of alternate data for analysis of credit scoring and predictive one:** There is large count of SMEs and financially barred individuals that are unable to use bank credit in cases when there is very little or no credit account history. So, giving a loan to such customers becomes very challenging for banks. However, now loans can be sanctioned with the usage of AI by fintech startups. They can gather and process data such as educational background, social media, police records, employment history, age, location, spending habits, and other things. With predictive analysis (using AI), one can compute the credit score, avoid bad loans, and give insight into the current demand of a client's credit and the next.

Now, several FinTech companies are using AI power-driven algorithms that are disturbing the loan industry with their AI solutions that increase markets [15, 25].

Regulatory Compliance, Fraud Prevention, and Detection and Prevention of Money Laundering: After the financial calamity in 2009, there was stress of compliance and risk management on the banks and financial services firms. Basel Accords provided risk management and unwieldy capital sufficiency compliance I, II, and II together with AML and KYC processes that are required for managing all types of risks in the system. It is also required when banks are using the credit system and various practices that can lead to fraud. Although this process requires time in the units of person hours due to huge paperwork, it opens the way for the usage of AI. If AI is implemented in this work, then this work can be completed within seconds by recognizing the patterns and reading the data in very less time. Using JP Morgan's COIN, millions of hours of work can be completed in minutes. The doubt is indicated and prediction of human activities done by anti-fraud-based AI products that mark variance. Further, for image recognition at ATMs, deep learning like AI techniques can be used along with real-time camera images that help expose and avoid crimes and fraud [17, 20].

## **4 Role of AI in the Finance Sector**

This section presents the contribution of various researchers in finance that have used AI algorithms to boost performance in the respective domain. Various Artificial Intelligence, deep learning applications in finance are presented in given sub-sections that have used Support vector machine (SVM), Recurrent Long Short Term Memory (LSTM), Backpropagation neural network (BPNN), Particle swarm optimization (PSO), DBN, K-nearest neighbour (KNN), Naïve Bayes, decision tree, Convolution neural network (CNN) and hybrid of one or two algorithms in financial distress, predicting credit card risk, Sentiment analysis, algorithm trading and stock price description that all comes under finance sector.

### ***4.1 Financial Distress in Finance Sector Using Artificial Intelligence***

For both practitioners and scholars, a great interest has been found by financial distress forecasting. The probability of financial distress can be estimated using a number of AI and statistical approaches. In this case, the prediction of financial distress for a system having a probability greater than the cutoff value is considered. This was improved by Bae [29] by collecting the financial data of MNCs located in the Korea Credit Guarantee Fund (KODIT) annually for the prediction of accuracy of financial distress problems [29].

Further, they have developed a radial basis function SVM (RSVM) based financial distress prediction model. For justification, they have compared their proposed RSVM with AI techniques and suggested a better financial distress predicting model that helps a chief finance officer make better decisions incorporating financial distress. Then, Hsieh et al. [30], proposed the SVM method and examined its predictive ability. Their proposed method uses the characteristics of a penalty function for generating predictions in a better way [30]. Further, presented an evolutionary artificial bee colony (EABC) algorithm for including the properties of Particle Swarm Optimization (PSO) in which a velocity and flying direction is given by each bee that optimize their proposed penalty guided SVM (PGSVM). For public industrial firms in Taiwan, an EABC-PGSVM was used for constructing a reliable prediction model and compared the proposed EABC-PGSVM with backpropagation neural network (BPNN), PGSVM optimized by the ABC algorithm (BPGSVM), and classic SVM optimized by the ABC algorithm (BSVM). Indifference to existing methods, Lin et al. [31], have proposed an approach for feature selection of FDP problems that combines expert knowledge with the wrapper approach [31]. Based on experts' domain knowledge, the financial features are categorized into seven classes. Fengyi has applied the wrapper method to search subsets of good features containing top candidates from each feature class. They have compared various scholarly models for concept verification that lead to feature selection methods. Their experiment indicates that the proposed method has selected feature set-based prediction model that gives a better outcome in terms of prediction accuracy than standard feature selection-based models. Yu-Pei Huang et al. [32], have also reviewed work done in predicting financial distress using ML algorithms [32]. Among all four supervised algorithms, the XGBoost gives a more accurate outcome in terms of FD prediction. The hybrid of the DBN-SVM model gives more accurate forecasts than using SVM or DBN classifiers individually.

Furthermore, a novel meta FDP framework was proposed by Wang et al. [33], it consists of feature regularising modules for the identification of discriminatory predictive power of number of features and enhances the aggregation over base classifiers using a probabilistic fusion module [33]. The results obtained from it show that the proposed RS2\_ER method can give an effective prediction on FDP. Then Sun et al. [34], have focused on effectively constructing class-imbalanced data streams based on dynamic FDP models [34]. The combination of SMOTE and AdaBoost SVM ensemble integrated with time weighting (ADASVM-TW) in which SMOTE stands for synthetic minority oversampling technique was utilised by them. In case of SMOTE, the class of every data batch is balanced before applying another approach for prediction modelling of dynamic financial distress. In the second one, the SMOTE is embedded with ADASVM-TW and designs a new sample weighting mechanism. For testing purposes, financial data of 2628 Chinese listed companies' dataset has been used that show both simple and embedding integration model that are able to improve the recognition ability for minority financial distress samples.

Numerous models are proposed for the detection of an occurrence of significant events in financial systems, but there is a need to automate significant events. In this concern, Rönqvist et al. [35], have presented a method for recognizing relevant

text and extracting natural language descriptions of events using deep learning [35]. Their model is leveraged by semantic vector representations unsupervised learning and supervised by entity names and dates a small set of event information on extensive test data. They have demonstrated the applicability of their news-based financial risk mainly related to bank distress and government interventions. Along with this, various researchers have proposed models for predicting contractor financial crises. So, Choi et al. [36], have proposed ensemble models based on voting that predict the financial distress of contractors for 2 or 3 years ahead of the prediction point using a financial distress definition based on finance [36]. The South Korean financial contractors' statements for the period 2007 to 2012 were used to evaluate the proposed model's performance. The results obtained show that the 0.940 and 0.910 receivers operating characteristic curve (AUC) values predict financial distress for each prediction year. In 2020, Said MARSO et al. [37] did an analysis on the performance of advanced cuckoo algorithm in terms of getting optimum weight of feedforward neural network and further named it CSFNN [37].

Further, to investigate the efficiency, they have compared the CSFNN with back-propagation feedforward neural network (BPNN) and Logistic Regression (LR) by applying it to two different periods of manufacturing sector collected data. The outcome in case of one year before bankruptcy gives an accuracy of 90.30% in the case of the CSFNN model and 88.33% and 82.15% in the case of the BPNN and LR model. It was 82.79% for CSFNN and 81.05% and 73.27% for BPNN and LR, respectively, in the case of three years before the bankruptcy.

#### ***4.2 Prediction of Credit Card Risk in the Finance Sector Using Artificial Intelligence***

In the analysis of credit card, a predictive performance of a broad class of binary classifiers was examined by Jones et al. [38], in which they used large sample of global credit ratings for the period of 1983–2013 [38]. The study discovered that the new classifiers outperform existing ones in cross-sectional and longitudinal test samples and are robust to a variety of data types and assumptions. They have concluded that simple classifiers can be used in more sophisticated approaches, mainly in having the main objective of interpretability of modelling exercise. The study has shown that financial credit scoring is crucial in the finance industry sector for assessing individuals' creditworthiness and enterprises. For performing this task, various statistics-based ML techniques have been employed, but in ML techniques, one of the significant challenges is the curse of dimensionality. So, to improve classification Jadhav et al. [39], have investigated feature selection in credit scoring problems and proposed a novel approach. Information Gain Directed Feature Selection algorithm (IGDFS) for it [39]. The proposed approach performs the feature ranking based on the information gain and GA wrapper (GAW) algorithm for propagating the top  $m$  features and then classifies it using SVM, KNN, and Naïve Bayes ML algorithms. The outcome



achieved inaccuracy shows that SVM gives the best results from the other two ML algorithms. This shows that SVM models give good outcome in credit scoring performance so, Tian, et al. [40], have also proposed state-of-the-art kernel-free fuzzy quadratic surface SVM model approach [40]. The results show that their proposed method performs well in classification and handling searching proper kernel function and complex model issues related to classical SVM models. Further, Masmoudi et al. [41] have modelled the payment default of loan subscribers using a discrete Bayesian network that includes a built-in clustering feature with latent variables [41]. The model was calibrated on loan contracts describing the actual dataset, and results obtained from it highlight a regime-switching of a default probability distribution. To deduce various researchers have employed the possible repayment behaviour of rejected credit applicants ML and statistical methods. Shen et al. [42], have used unsupervised transfer learning and 3-way decision theory to propose a novel 3-stage reject inference learning framework [42]. The usability of the proposed framework shows its applicability in rejecting inference and handling adverse transfer learning problems. For validation of the proposed framework, Chinese credit data was considered, and the outcome obtained from it shows the superiority in credit risk management applications. Further, Wang et al. [43], have focused on a comparative assessment of five popular classifiers' performance in credit scoring [43]. The classifiers considered were LR, Random Forest, KNN, Naïve Bayes and Decision Tree. The study found that all classifiers have their pros and cons, so saying which one is best is emphatic. However, in terms of AUC, accuracy, recall, and precision, a better outcome is achieved using Random Forest.

Furthermore, for the credit scoring problem, the suitability of dynamic selection techniques was evaluated by Leopoldo Melo [44]. They have also presented Reduced Minority k-Nearest Neighbors (RMkNN) to enhance the state-of-the-art dynamic selection techniques in local regions for imbalanced credit scoring datasets. As compared to state-of-the-art, better prediction performance is achieved using the proposed technique. The other main benefit of RMkNN is that there is no need for any sampling or pre-processing method for generating a dynamic selection dataset (called DSEL). For predicting whether the loan will be repaid in the P2P platform or not, a benchmarking study of various credit risk scoring models was proposed by Vincenzo [45]. For analysis of the experiment, an 877,956 samples real social lending platform (Lending Club) dataset was used and evaluated results in terms of specificity, AUC, and Sensitivity. In the end, the best three approaches have been evaluated using various eXplainable Artificial Intelligence (XAI) tools.

For the prediction of bankruptcy and finance activity, credit risk assessment is a critical task that has been explored using ML and statistical methods. To further enhance the performance of credit modelling use of ensemble strategies has been suggested in recent works. So, Florez-Lopez et al. [28], have explored various complementary sources of diversity for optimizing the model's structure that leads to a manageable number of decision rules without affecting the performance [28]. The empirical results suggest that CADF is a good solution compared to individual classifiers and RF, gradient boosting-like ensemble strategies for credit risk problems. By seeing the improvement using AI/DL approaches, Huang et al. [46], have

also used a probabilistic neural network (PNN) that gives a minimum error rate and the second type of error and the highest AUC value [46].

Further, the fraud detection problem was phrased as LSTM and sequence classification task to incorporate transaction sequences. Jurgovsky et al. [47], have also integrated the traditional feature aggregation approach and reported results in traditional retrieval metrics [47]. Comparing the proposed algorithm with the baseline RF classifier shows an improvement in detection accuracy on offline transactions. From manual features, aggregation approaches benefits are achieved by both sequential and non-sequential learning approaches.

To construct a credit risk assessment model Zhang et al. [48], have proposed a new approach for peer to peer lending market [48]. They first used a Transformer encoder to extract the textual features from the loan description then combined them with loan application-derived challenging features and final loan features. Then send the combined features into a two-layer Feedforward NN for predicting the probability of default loans. The proposed approach was tested on Renrendai loan data from the Chinese market and LendingClub loan data from the American market datasets. The results show that a better outcome is achieved by the model in which the textual loan description is considered compared to loan default prediction, and the best outcome is achieved under AUC and G-mean metrics.

Further, Golbayani et al. [49], did a survey and gave a comparative outcome of results obtained by various ML and AI techniques in predicting credit rating [49]. Then they applied RF, SVM, Multilayer perceptron, and Bagged decision tree techniques on the same datasets and evaluated the results using 10 tenfold cross-validation techniques. The results show the best performance is achieved using a Decision tree-based model.

### ***4.3 Sentiment Analysis in the Finance Sector Using Artificial Intelligence***

The importance of analysing the massive volume of text from social networks and websites has been raised by developing online virtual communities. To develop a public mood dynamic prediction model has been developed by Chen et al. [50], by analysing online news articles and financial blogs [50]. This has been done concerning behavioural finance perspectives and characteristics of online financial communities. To Taiwan sentiment analysis investors opinion mining and big data approaches are applied in their work and verified their proposed model using China-Time.com, Google stock market news, Yahoo stock market news, and cnYES.com experimental datasets. The results obtained from it show that big data analysis techniques for assessing the emotional content of commentary on financial and current stock issues can be forecasted effectively. Further, Twitter data 1 2 has been considered by company stock prices and served the need for scoring the impression carried

out for a particular firm. Das et al. [51], have made a classifying model from historical data that can improve outcomes [51]. For humongous data processing, spark streaming has been considered, along with Apache Flume and Twitter API-like data ingestion tools used for further implementation and analysis. Xu et al. [52], have also presented a continuous naïve Bayes learning framework to review multi-domain and large-scale e-commerce platform product sentiment classification [52]. They have also extended the naïve Bayes parameter estimation mechanism to a continuous learning style and then proposed various ways of fine-tuning the learned distribution based on three types of assumptions for adapting better to different domains. The experiment was conducted on the movie reviews and the Amazon product sentiment dataset.

Various news articles are engaged in prediction processes, but combining the technical indicators from news and stock price sentiments and making prediction models learn sequential information within time series better is still of concern. So, Li et al. [53], have proposed a new stock predicting system that is able to represent numerical price data using technical indicators and analysis and further represent the textual news articles using sentiment vectors of sentiment analysis [53]. Further, they have set up the deep learning model for learning the sequential information within the series of market snapshots that are constructed by news sentiments and technical indicators.

Rich source of information is represented by textual materials that improve the decision making of organizations, businesses, and people. Pröllochs et al. [54], have proposed an approach in which they have taken document level labels as input and then learn a document level labels-based negation policy [54]. There are various limitations in existing models. To address these problems, Mohammad Ehsan Basiri et al. [55], have proposed an attention-based Bidirectional CNN-RNN Deep Model (ABCDM) [55]. In this, they have used two independent bidirectional GRU, and ABCDM extracts LSTM layers and both past and future contexts by considering the flow of temporal information in both directions. Further, to put more or less emphasis on various words an ABCDM bi-direction layer was used that reduces the dimension of feature and extract the position invariant local features. On sentiment polarity detection, ABCDM effectiveness detection is considered the most common and necessary sentiment analysis task. For experiment purposes, three Twitter and five review datasets are used, and ABCDM results are compared with six proposed DNNs for sentiment analysis.

#### ***4.4 Algorithm Trading for Finance Traders Using Artificial Intelligence***

The motive of Rys et al. [56], analysed and formulated the machine learning approaches with fixed strategy optimization specificity parameters [56]. Sensitivity performance is the most critical problem for little change in parameter and number of local extrema distribution over the solution space in a distinctive way. The approaches were designed for significant shortening of computation time without affecting the substantial strategy quality of loss. Their method was operated on 20 years of daily price sample data and presented three sets of two asset portfolios. The strategy was traded on DAX and SPX index futures in the case of first case, and in the second case, it was done on MSFT and AAPL stocks, and then the final case was done on CBF and HGF commodities futures.

For financial markets trading like forex and stock, AI has been increased, and out of all, reinforcement learning has become prevalent for financial traders. Meng et al. [57], have reviewed all current forex/stock predictions in which reinforcement learning has been used as a direct ML approach [57]. All the articles reviewed in this work have various unrealistic assumptions like no bid or ask spread issues and no liquidity and transaction costs. On reinforcement learning algorithms profitability, a significant impact has been seen in transaction costs compared to baseline-tested algorithms. They have also given a performance comparison between reinforcement learning and other DL or ML models and assessed the impact of bid/ask spread on the profitability of transaction costs. From overall work, it has been found that reinforcement learning in forex or stock trading helps in early development and also stated that there is a need for a reliable approach in the same domain. Then, Fengqian et al. [58], have used real-time financial data and processed it using K-line theory and candlesticks as a generalization price movement for a period that helps in de-noising [58].

Further, a decomposition of candlesticks is done into various subparts by using a specified Spatio-temporal relationship based on which subparts cluster analysis was obtained for getting the features of learning. Along with this, K-lines are used to clustered learning features that are added into the model, and unknown environment adaptive control parameters are realized using a deep reinforcement learning approach for realizing the high-frequency transaction strategy. To verify model performance, they have used various financial derivatives transactions like financial features, commodity features, and stocks. They also compared the proposed approach with fuzzified price, K-lines, and price-based methods. Fuzzy neural networks and recurrent neural network-like prediction-based approaches are used to verify the proposed method's accuracy which shows a higher prediction accuracy and robustness of the proposed method.

## 4.5 Prediction of Stock Price Indexing

There are various problems associated with the prediction of direction of movement of the stock and stock price index for Indian stock markets. Jigar Patel et al. [59], have compared SVM, NB, RF, and ANN four prediction models with two approaches for input to these models [59]. Computation of ten technical parameters computation is involved with the first approach for input data in which they have used open, low, close, and high prices stock trading data. On the other hand, these technical parameters are represented as trend deterministic data in the second approach. Then for these two input approaches, prediction model accuracy was evaluated for which ten years of historical data was used that is taken from 2003 to 2012 of Infosys and Reliance industries stocks. The results obtained from it show that out of the other three prediction models, the random forest can give better performance for the first approach of input data. Further, Bisoi et al. [60], have focused on two objectives, namely daily trend prediction and day ahead stock price prediction using a Robust Kernel-based Extreme Learning Machine (RKELM) integrated with VMD in which Differential evolution algorithm was used for the optimization of kernel function parameters DE-VMD-RKELM [60]. In the end, trend prediction was compared with SVM, ANN, and Naïve Bayes classifier that shows the superiority of the proposed model over other predictive methods.

With the use of AI and an increase in computational capabilities programmed prediction approaches are proven to be more efficient when used in the prediction of stock prices. For the prediction of next day closing of 5 different sectors of companies, ANN and RF techniques were used by Vijn et al. [61]. New variables are created by close, low, high, and open stock prices as financial data used as the input to the model. MAPE and RMSE standard strategic indicators are used to evaluate the models.

Solutions to various challenging problems have been provided by the current flow in the research of DL. New methods for these problems have been adopted in quantitative analysis. However, due to non-stationary financial data-like problems, significant challenges must be overcome before using DL. So, Tsantekidis et al. [62], have proposed a new approach for constructing stationary features that allow DL models to be applied efficiently [62]. The tasks of mid-price limited order book movements task are used for thoroughly testing these features. They have evaluated Convolutional Neural networks (CNN) and Recurrent Long Short Term Memory (LSTM) networks like DL models. The author has evaluated the novel model in which LSTM and CNN useful features' ability are extracted for analysing the time series. The outcome achieved from it shows that the combined model gives better results than tested individual CNN and LSTM models in the prediction horizons. Then, Chalvatzis et al. [63], have used tree-based models to test their proposed approach along with deep LSTM neural networks [63]. The results obtained on testing methods for the period of 2010–2019 show a 350, 403, 497 and 333% of the overall model achieved cumulative returns using S&P 500, Dow Jones Industrial Average (DJIA), NASDAQ, and Russell 2000 stock indices, respectively.

Furthermore, Hulu et al. [64], have proposed a novel stock closing price forecasting framework with higher prediction than traditional models (Hulu et al. 2020). This deep hybrid framework contains the predictor optimization method, deep learning predictor part, and data processing part components. In this pre-processing is done using empirical wavelet transform and in data processing a post processing is done using an outlier robust extreme learning machine. The primary part of this composite frame is an LSTM network-based deep learning predictor network that is jointly optimized by the PSO algorithm and dropout strategy. In their hybrid framework, every algorithm plays its functions for getting better prediction accuracy [64]. For forecasting experiments, three challenging datasets are used to verify the performance of their proposed model, and used various comparative models to prove the proposed framework's effectiveness.

## 5 Comparative Analysis

This section shows the comparative analysis of learning technologies applied in finance applications and classifiers employed by researchers for a particular application and their results. The data presented below in Table 2 will help understand which machine learning technology is better to implement in which finance application. This table can help in making Hybrid Models for improvising results in various aspects of finance applications.

Finance is all about data, and we have seen handling such big data and processing it. Both are challenging tasks [81–83]. Machine learning or AI is a process of learning that is training from the data and making predictions on testing based on the training data, which means training is a crucial step in the machine learning process. As data is big data, it makes the processing speed low. AI is an expensive technology [84, 85]. As we do not have a universal AI algorithm that can be applied to most finance applications, every time, a new application requires a new version of algorithms. It makes it very costly. As implementation cost is high, sometimes this technique fails to give accurate results and gives results with a high false rate. Acceleration in the access of databases or financial data is vital because accuracy somehow depends on it. So it can be suggested that this survey chapter can help researchers to have a brief overview of all available machine learning techniques applied in finance, and this can help them to build a hybrid model which can, in return, help in improvisation of accuracy in terms of processing time, cost and as well as performance.

**Table 2** Comparative analysis of applications and technologies implemented

Application	Author	Learning technique	Dataset used	Classifiers	Parameters
Predicting stock price index	Kyoung-jae and Han [65]	Supervised learning	Korea	Neural networks, GA	Accuracy = 61.70%
	Kara et al. [66]	Supervised learning	Turkey	Neural networks, SVM	Accuracy = 75.74%
	Chiang et al. [67]	Supervised learning	Multiple	Neural networks	Accuracy = 87.80%
	Patel et al. [68]	Supervised learning	India	Neural networks, SVM, RF,NB	ANN= 86.69%, SVM = 89.33%, RF=89.98% NB = 90.19%
	Wang [69]	Supervised learning	Taiwan	Fuzzy logic	Accuracy= 93%
Financial distress prediction model	Chen et al. [70]	Supervised learning	Taiwan stock exchange corporation	ANN and DM	Accuracy = 82.14%
Algorithmic trading rule	Rodríguez et al. [71]	Supervised learning	Spain	Neural networks	Accuracy = 54–58%
	Giacomet et al. [72]	Supervised learning	North American and Brazilian stock markets	Ensemble of neural networks	Capital hit rate = 60% And capital Gain = 56% in 166 days
	Taylor et al. [73]	Supervised learning	S&P 500	Artificial neural network	Accuracy = 74.45%
Fraud detection	Maes et al. [74]	Supervised learning	Europay International (EPI)	ANN & BBN	Accuracy of ANN = 60% BNN = 68% (±10% False POS)
	Raj et al. [75]	Supervised learning	FDS	Fuzzy logic	Accuracy = 100%
	Aleskerov et al. [76]	Supervised learning	Tested on synthetically generated data	Neural network based data mining	Accuracy = 85%
Credit scoring	Sayjadah et al. [77]	Supervised learning	Real-time data	Random forest	Accuracy = 81.81%

(continued)

**Table 2** (continued)

Application	Author	Learning technique	Dataset used	Classifiers	Parameters
	Bing Zhu et al. [78]	Supervised learning	Real-world dataset from a Chinese consumer finance company	Relief algorithm + CNN	91.64%
Chatbots	Muangkammuen et al. [79]	Supervised learning	Thai text	RNN	Question recognition accuracy = 86.36% Answering accuracy = 93.20%
Portfolio management	Jiang et al. [23]	Supervised learning	Poloniex	Deep reinforcement learning method	tenfold returns in 1.8 months
High-frequency trading	Yang et al. [80]	Supervised learning	E-Mini S&P 500 future market	Inverse Reinforcement learning	Accuracy = 90%

## 6 Conclusion

In the presented finance chapter, we have seen that machine learning or AI is a good subset of data science, rapidly undergoing development. The financial market is exceptionally well suited for it, and its potential applications in finance are constantly growing. With the introduction of ML and AI in financial systems, number of data can be analysed, store, interpreted and calculated without explicit programming. This chapter addressed machine learning, artificial intelligence techniques and briefly commented on popular models, such as ANN, SVM, CNN, BFNN and RF, and presented a systematic review of various terms related to machine learning and artificial intelligence techniques that concern the finance sector. This provides an abstract view to the reader of artificial intelligence and its usage in various domains of the financial domain—the benefits and issues related to implementing digital solutions in the financial system. From the chapter, we also tried to present the earlier work done in this research field and have done comparative analysis that helps other researchers to use it efficiently and appropriately in future developments.



## References

1. Aziz S, Dowling M (2018) Machine learning and AI for risk management. Springer International Publishing, pp 33–50
2. Prado MLD (2019) Ten applications of financial machine learning. SSRN Electron J, 1–16
3. Chen K, Yadav A, Khan A, Zhu K (2020) Credit fraud detection based on hybrid credit scoring model. International conference on computational intelligence and data science (ICCIDS 2019), 2–8
4. Buchanan BG (2019) Artificial intelligence in finance. Work was supported by The Alan Turing Institute
5. Klein KJ, Knight AP (2015) Innovation Implementation: overcoming the challenge. *Curr Dir Psychol Sci* 14:243–246
6. Mukherjee S, Shaw R (2016) Big data—concepts, applications, challenges and future scope. *Int J Adv Res Comput Commun Eng* 5:66–74
7. Henrique BM, Sobreiro VA, Kimura H (2019) Literature review: machine learning techniques applied to financial market prediction R. *Expert Syst Appl* 124:226–251
8. Heaton JB, Polson NG (2016) Deep learning in finance. *Appl Stochastic Models Bus Industry*, 1–20
9. Provost F, Fawcett T (2013) Data science and its relationship data-driven decision making. Foster provost and tom fawcett. *Big Data* 1:51–59
10. Marrara S, Azzini S (2018) Opportunities and risks for data science in organizations: banking, finance, policy, special session overview, 612–613
11. Yang B (2019) Construction of logistics financial security risk ontology model based on risk association and machine learning. *Saf Sci* 123:1–11
12. Zhu Y, Zhou L, Xie C, Wang G, Nguyen TV (2019) Forecasting SMEs credit risk in supply chain finance with an enhanced hybrid ensemble machine learning approach. *Int J Prod Econ* 211:22–33
13. Fan Y (2015) Domain ontology construction method and empirical research. Wuhan University Press, Wuhan, pp 39–40
14. Florez-lopez R (2007) Modelling of insurers' rating determinants. An application of machine learning techniques and statistical models. *Europ J Oper Res* 183:1488–1512
15. Dorransoro R, Ginel F, Carmen S, Cruz CS (1997) Neural fraud detection in credit card operations. *IEEE Trans Neural Network* 8:827–834
16. Wang C, Han D, Luo S (2019) A deep learning approach for credit scoring of peer-to-peer lending using attention mechanism LSTM. *IEEE Access* 7:2161–2168
17. Duhart B, Hernández-gress N (2016) Review of the principal indicators and data science techniques used for the detection of financial fraud and money laundering. International conference on computational science and computational intelligence (CSCI), 1397–1398
18. Harvey C, Liu Y (2015) Backtesting. *J Portf Manag* 42:13–28
19. Agarwal P (2019) Redefining banking and financial industry through the application of computational intelligence. 2019 Advance science engineering technology international conference, pp 1–5
20. Mehrotra A (2019) Artificial intelligence in financial services—need to blend automation with human touch. 2019 International conference automobile computer technology management, pp 342–347
21. Zhang P, Yu K, Yu JJ, Khan SU, Member S (2017) QuantCloud: big data infrastructure for quantitative finance on the cloud. *IEEE Trans Big Data* 4:368–380
22. Lin W, Hu Y, Tsai C (2012) Machine learning in financial crisis prediction: a survey. *IEEE Trans Syst Man Cybern. Part C (Appl Rev)* 42:421–436
23. Jiang Z (2017) Cryptocurrency portfolio management with deep reinforcement learning, IntelliSys
24. Cheng T (2018) AI Robo-advisor with big data analytics for financial services. 2018 IEEE/ACM international conference advance social networks analysis mining, pp 1027–1031

25. Guo S, He H, Huang X (2019) A multi-stage self-adaptive classifier ensemble model with application in credit scoring. *IEEE Access* 7:78549–78559
26. Ben-David, Arie, Frank E (2009) Accuracy of machine learning models versus “hand crafted” expert systems—a credit scoring case study. *Expert Syst Appl*, 5264–5271
27. Snihovyi, Oleksandr, Ivanov O, Kobets V (2018) Implementation of Robo-advisors using neural networks for different risk attitude investment decisions. 2018 International conference on intelligent systems (IS). *IEEE*
28. Florez-Lopez R, Ramon-Jeronimo JM (2015) Enhancing accuracy and interpretability of ensemble strategies in credit risk assessment, A correlated-adjusted decision forest proposal. *Expert Syst Appl*, 5737–5753
29. Bae JK (2012) Predicting financial distress of the South Korean manufacturing industries. *Expert Syst Appl*, 9159–9165
30. Hsieh TJ, Hsiao HF, Yeh WC (2012) Mining financial distress trend data using penalty guided support vector machines based on hybrid of particle swarm optimization and artificial bee colony algorithm. *Neurocomputing*, 196–206
31. Lin F, Liang D, Yeh CC, Huang JC (2014) Novel feature selection methods to financial distress prediction. *Expert Syst Appl*, 2472–2483
32. Huang YP, Yen MF (2019) A new perspective of performance comparison among machine learning algorithms for financial distress prediction. *Appl Soft Comput J*, 1–14
33. Wang G, Ma J, Chen G, Yang Y (2020) Financial distress prediction: regularized sparse-based Random Subspace with ER aggregation rule incorporating textual disclosures. *Applied Soft Comput J*, 1–17
34. Sun J, Li H, Fujita H, Fu B, Ai W (2020) Class-imbalanced dynamic financial distress prediction based on Adaboost-SVM ensemble combined with SMOTE and time weighting. *Inf Fusion*, 128–144
35. Rönnqvist S, Sarlin P (2016) Bank distress in the news: describing events through deep learning. *Neurocomputing*, 1–32
36. Choi H, Son H, Kim C (2018) Predicting financial distress of contractors in the construction industry using ensemble learning. *Expert Syst Appl*, 1–44
37. Marso S, Merouani ME (2020) Predicting financial distress using hybrid feedforward neural network with cuckoo search algorithm. *International workshop on statistical methods and artificial intelligence (IWSMAI 2020)*, pp 1134–1140
38. Jones S, Johnstone D, Wilson R (2015) An empirical evaluation of the performance of binary classifiers in the prediction of credit ratings changes. *J Banking Finance*, 1–43
39. Jadhav S, He H, Jenkins K (2018) Information gain directed genetic algorithm wrapper feature selection for credit rating. *Appl Soft Comput*, 1–50
40. Tian Y, Yong Z, Luo J (2018) A new approach for reject inference in credit scoring using kernel-free fuzzy quadratic surface support vector machines. *Appl Soft Comput J*, 1–26
41. Masmoudi K, Abid L, Masmoudi A (2019) Credit risk modeling using Bayesian network with a latent variable. *Expert Syst Appl*, 157–166
42. Xu F, Pan Z, Xia R (2020) E-commerce product review sentiment classification based on a naive Bayes continuous learning framework. *Information processing and management*, pp 1–7
43. Wang Y, Zhang Y, Lu Y, Yu X (2020) A Comparative Assessment of Credit Risk Model based on Machine Learnin: a case study of bank loan data. *International conference on identification, information and knowledge in the internet of things*, pp 141–149
44. Junior LM, Nardini FM, Renso C, Trani R, Macedo JA (2020) A novel approach to define the local region of dynamic selection techniques in imbalanced credit scoring problems. *Expert Syst Appl*, 1–16
45. Moscato V, Picariello A, Sperlí G (2021) A benchmark of machine learning approaches for credit score prediction. *Expert Syst Appl*, 1–8
46. Huang X, Liu X, Ren Y (2018) Enterprise credit risk evaluation based on neural network algorithm. *Cognit Syst Res*, 1–11
47. Jurgovsky J, Granitzer M, Ziegler K, Calabretto S, Portier PE, He-Guelton L, Caelen O (2018) Sequence classification for credit-card fraud detection. *Expert Syst Appl*, 1–30

48. Zhang W, Wang C, Zhang Y, Wang J (2020) Credit risk evaluation model with textual features from loan descriptions for P2P lending. *Electronic commerce research and applications*, pp 1–31
49. Golbayani P, Florescu I, Chatterjee R (2020) Comparative study of forecasting corporate credit ratings using neural networks, support vector machines, and decision trees. *North Amer J Econ Finance*, 1–16
50. Chen MY, Chen TH (2017) Modeling public mood and emotion: Blog and news sentiment and socio-economic phenomena. *Future Gen Comput Syst*, 1–27
51. Das S, Behera RK, kumar M, Rath SK (2018) Real-time sentiment analysis of twitter streaming data for stock prediction. *Procedia Comput Sci*, 956–964
52. Shen F, Zhao X, Kou G (2020) Three-stage reject inference learning framework for credit scoring using unsupervised transfer learning and three-way decision theory. *Decision Support Syst*, 1–10
53. Li X, Wu P, Wang W (2019) Incorporating stock prices and news sentiments for stock market prediction: a case of Hong Kong. *Inf Process Manag*, 1–19
54. Pröllochs N, Feuerriegel S, Lutz B, Neumann D (2020) Negation scope detection for sentiment analysis: A reinforcement learning framework for replicating human interpretations. *Inf Sci*, 205–221
55. Basiri ME, Nemati S, Abdar M, Cambria E, Acharya UR (2021) ABCDM: an attention-based bidirectional CNN-RNN deep model for sentiment analysis. *Future Gen Comput Syst*, 279–294
56. Rys P, Slepaczuk R (2019) Machine learning in algorithmic trading strategy optimization—implementation and efficiency. *Central Europ Econ J*, 206–229
57. Meng TL, Khushi M (2019) Reinforcement learning in financial markets. *MDPI*, 1–17
58. Fengqian D, Chao L (2020) An adaptive financial trading system using deep reinforcement learning with candlestick decomposing features. *IEEE Access*, 63666–63678
59. Patel J, Shah S, Thakkar P, Kotecha K (2014) Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert Syst Appl*, 1–14
60. Bisoï R, Dash PK, Parida AK (2018) Hybrid variational mode decomposition and evolutionary robust kernel extreme learning machine for stock price and movement prediction on daily basis. *Appl Soft Comput J*, 1–66
61. Vijh M, Chandola D, Tikkiwal VA, Kumar A (2020) Stock closing price prediction using machine learning techniques. *International conference on computational intelligence and data science (ICCIDIS)*, pp 599–606
62. Tsantekidis A, Passalis N, Tefas A, Kannianen J, Gabbouj M, Iosifidis A (2020) Using deep learning for price prediction by exploiting stationary limit order book features. *Appl Soft Comput J*, 1–10
63. Chalvatzis C, Hristu-Varsakelis D (2020) High-performance stock index trading via neural networks and trees. *Appl Soft Comput J*, 1–22
64. Carta S, Ferreira A, Podda AS, Recupero DR, Sanna A (2020) Multi-DQN: an ensemble of deep Q-learning agents for stock market forecasting. *Expert Syst Appl*, 1–46
65. Kyoung-jae K, Han I (2000) Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. *Expert Syst Appl* 19:125–132
66. Yakup K, Boyacioglu MA, Baykan OK (2011) Predicting direction of stock price index movement using artificial neural networks and support vector machines: the sample of the Istanbul Stock Exchange. *Expert Syst Appl* 38:5311–5319
67. Chiang WC, Enke D, Wu T, Wang R (2016) An adaptive stock index trading decision support system. *Expert Syst Appl* 59:195–207
68. Patel J, Shah S, Thakkar P, Kotecha IK (2015) Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert Syst Appl* 42:259–268
69. Wang YF (2003) Mining stock price using fuzzy rough set system. *Expert Syst Appl* 24:13–23
70. Chen W-S (2009) Du YK. Using neural networks and data mining techniques for the financial distress prediction model. *Expert Syst Appl* 36:4075–4086

71. Fernandez-Rodriguez F, Gonzalez-Martel C, Sosvilla-Rivero S (2000) On the profitability of technical trading rules based on artificial neural networks: Evidence from the Madrid stock market. *Econ Lett* 69:89–94
72. Giacomel F, Galante R, Pereira A (2015) An algorithmic trading agent based on a neural network ensemble: A case of study in North American and Brazilian stock markets. *IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology (WI-IAT)*, 2
73. Taylor B, Kim M, Choi A (2014) Automated stock trading algorithm using neural networks. *Proceedings of the 2nd international conference on intelligent technologies and engineering systems (ICITES2013)*. Springer, Cham, pp 849–857
74. Maes S, Tuyls K (2002) Credit card fraud detection using Bayesian and neural networks. *Proceedings of the 1st international naiso congress on neuro fuzzy technologies*, 1–8
75. Raj BES, Portia AA (2011) Analysis on credit card fraud detection methods. *International conference on computer, communication and electrical technology (ICCCET)*, IEEE, pp 152–156
76. Aleskerov E, Freisleben B, Rao B (2002) Cardwatch: A neural network based database mining system for credit card fraud detection. *Proceedings of the IEEE/IAFE computational intelligence for financial engineering (CIFEr)*. IEEE, pp 220–226
77. Yashna S, Hashem IAT, Alotaibi F, Kasmiran KA (2018) Credit card default prediction using machine learning techniques. *2018 Fourth international conference on advances in computing, communication & automation (ICACCA)*, IEEE, pp 1–4
78. Zhu B, Yang, W, Wang H, Yuan Y (2018) A hybrid deep learning model for consumer credit scoring. *International conference on artificial intelligence and big data (ICAIBD)*, IEEE, pp 205–208
79. Muangkammuen P, Intiruk N, Saikaew KR (2018) Automated Thai-FAQ Chatbot using RNN-LSTM. *22nd International computer science and engineering conference (ICSEC)*, IEEE, pp 1–4
80. Steve Y, Paddrik M, Hayes R, Todd A, Kirilenko A, Beling P, Scherer W (2012) Behavior based learning in identifying high frequency trading strategies. *IEEE Conference on computational intelligence for financial engineering & economics (CIFEr)*, pp 1–8
81. Unadkat V (2018) Deep learning for financial prediction. *2018 International conference circuits system digital enterprise technology*, pp 1–6.
82. Gong XL, Liu XH, Xiong X, Zhang W (2019) Financial systemic risk measurement based on causal network connectedness analysis. *Int Rev Econ Financ* 64:290–307
83. Luo J, Xin Y, Tian Y (2020) Unsupervised quadratic surface support vector machine with application to credit risk assessment. *Europ J Oper Res*, 1008–1017
84. Randhawa K, Loo CK, Seera M, Lim CP, Nandi AK (2018) Credit card fraud detection using AdaBoost and majority voting. *IEEE Access* 6:14277–14284
85. Shenyong X, Yu H, Wu Y, Peng Z, Zhang Y (2017) Self-evolving trading strategy integrating internet of things and big data. *IEEE Internet Things J* 5:2518–2525