Machine learning in business and finance: a literature review and research opportunities

Hanyao Gao[1], Gang Kou[2*], Haiming Liang[1], Hengjie Zhang[3], Xiangrui Chao[1], Cong-Cong Li[5] and Yucheng Dong[1,4*]

*Correspondence:
kougang@swufe.edu.cn;
ycdong@scu.edu.cn

[1] Business School, Sichuan University, Chengdu 610065, China
[2] School of Business Administration, Faculty of Business Administration, Southwestern University of Finance and Economics, Chengdu 611130, China
[3] Business School, Hohai University, Nanjing 211100, China
[4] Xiangjiang Laboratory, Changsha 410205, China
[5] School of Economics and Management, Southwest Jiaotong University, Chengdu 610031, China

Abstract

This study provides a comprehensive review of machine learning (ML) applications in the fields of business and finance. First, it introduces the most commonly used ML techniques and explores their diverse applications in marketing, stock analysis, demand forecasting, and energy marketing. In particular, this review critically analyzes over 100 articles and reveals a strong inclination toward deep learning techniques, such as deep neural, convolutional neural, and recurrent neural networks, which have garnered immense popularity in financial contexts owing to their remarkable performance. This review shows that ML techniques, particularly deep learning, demonstrate substantial potential for enhancing business decision-making processes and achieving more accurate and efficient predictions of financial outcomes. In particular, ML techniques exhibit promising research prospects in cryptocurrencies, financial crime detection, and marketing, underscoring the extensive opportunities in these areas. However, some limitations regarding ML applications in the business and finance domains remain, including issues related to linguistic information processes, interpretability, data quality, generalization, and the oversights related to social networks and causal relationships. Thus, addressing these challenges is a promising avenue for future research.

Keywords: Machine learning, Business, Finance, Marketing

Introduction

The rapid development of information and database technologies, coupled with notable progress in data analysis methods and computer hardware, has led to an exponential increase in the application of ML techniques in various areas, including business and finance (Ghoddusi et al. 2019; Gogas and Papadimitriou 2021; Chen et al. 2022; Hoang and Wiegratz 2022; Nazareth and Ramana 2023; Ozbayoglu et al. 2020; Xiao and Ke 2021). The progress in ML techniques in business and finance applications, such as marketing, e-commerce, and energy, has been highly successful, yielding promising results (Athey and Imbens 2019). Compared to traditional econometric models, ML techniques can more effectively handle large amounts of structured and unstructured data, enabling rapid decision-making and forecasting. These benefits stem from ML techniques' ability to avoid making specific assumptions about the functional form, parameter distribution,

or variable interactions and instead focus on making accurate predictions about the dependent variables based on other variables.

Exploring scientific databases, such as the Thomson Reuters Web of Science, reveals a significant exponential increase in the utilization of ML in business and finance. Figure 1 illustrates the outcomes of an inquiry into fundamental ML applications in emerging business and financial domains over the past few decades. Numerous studies in this field have applied ML techniques to resolve business and financial problems. Table 1 lists some of their applications. Boughanmi and Ansari (2021) developed a multimodal ML framework that integrates different types of non-parametric data to accommodate diverse effects. Additionally, they combined multimedia data in creative product settings and applied their model to predict the success of musical albums and playlists. Zhu et al. (2021) asserted that accurate demand forecasting is critical for supply chain efficiency, especially for the pharmaceutical supply chain, owing to its unique characteristics. However, a lack of sufficient data has prevented forecasters from pursuing advanced models. Accordingly, they proposed a demand forecasting framework that "borrows" time-series data from many other products and trains the data with advanced ML models. Yan and Ouyang (2018) proposed a time-series prediction model that combines wavelet analysis with a long short-term memory neural network to capture the complex features of financial time series and showed that this neural network had a better prediction effect. Zhang et al. (2020a, b) employed a Bayesian learning model with a rich dataset to analyze
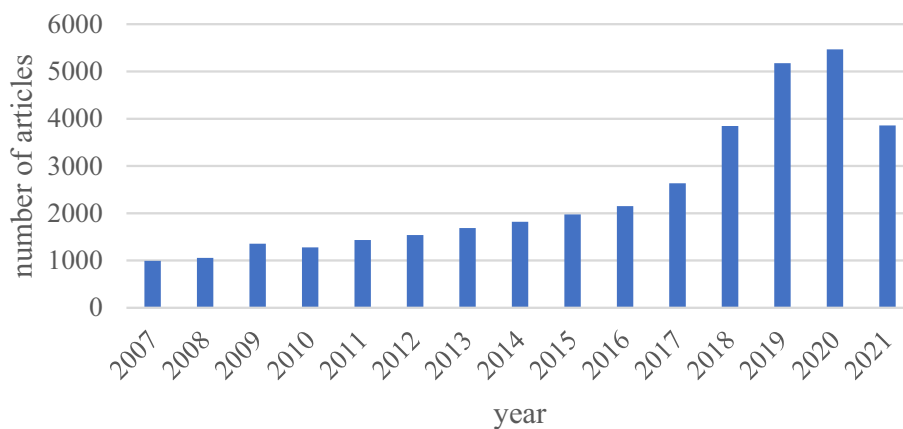

Fig. 1 Trend of articles on applied ML techniques in business and finance (2007–2021)

Table 1 ML techniques applied in the business and finance domains

References	Research issues	Machine learning techniques
Boughanmi and Ansari (2021)	Product forecasting	Bayesian networks
Zhu et al. (2021)	Supply-chain demand forecasting	Recurrent neural networks
Yan and Ouyang (2018)	Stock price prediction	Long short-term memory, support Vector machine, k-nearest neighbor
Zhang et al. (2020a, b)	Customer demand	Bayesian networks
Chen et al. (2017)	Credit scoring	Support vector machine
Lago et al. (2018)	Energy prices predicting	Deep neural networks

the decision-making behavior of taxi drivers in a large Asian city to understand the key factors that drive the supply side of urban mobility markets.

Several review papers have explored the potential of ML to enhance various domains, including agriculture (Raj et al. 2015; Coble et al. 2018; Kamilaris and Prenafeta-Boldu 2018; Storm et al. 2020), economic analysis (Einav and Levin 2014; Bajari et al. 2015; Grimmer 2015; Nguyen et al. 2020; Nosratabadi et al. 2020), and financial crisis prediction (Lin et al. 2012; Canhoto 2021; Dastile et al. 2020; Nanduri et al. 2020). Kou et al. (2019) conducted a survey encompassing research and methodologies related to the assessment and measurement of financial systemic risk that incorporated various ML techniques, including big data analysis, network analysis, and sentiment analysis. Meng and Khushi (2019) reviewed articles that focused on stock/forex prediction or trading, where reinforcement learning served as the primary ML method. Similarly, Nti et al. (2020) reviewed approximately 122 pertinent studies published in academic journals over an 11-year span, concentrating on the application of ML to stock market prediction.

Despite these valuable contributions, it is worth noting that the existing review papers primarily concentrate on specific issues within the realm of business and finance, such as the financial system and stock market. Consequently, although a substantial body of research exists in this area, a comprehensive and systematic review of the extensive applications of ML in various aspects of business and finance is lacking. In addition, existing review articles do not provide a comprehensive review of common ML techniques utilized in business and finance. To bridge the aforementioned gaps in the literature, we aim to provide an all-encompassing and methodological review of the extensive spectrum of ML applications in the business and finance domains. To begin with, we identify the most commonly utilized ML techniques in the business and finance domains. Then we introduce the fundamental ML concepts and frequently employed techniques and algorithms. Next, we systematically examine the extensive applications of ML in various sub-domains within business and finance, including marketing, stock markets, e-commerce, cryptocurrency, finance, accounting, credit risk management, and energy. We critically analyze the existing research that explores the implementation of ML techniques in business and finance to offer valuable insights to researchers, practitioners, and decision-makers, thereby facilitating better-informed decision-making and driving future research directions in this field.

The remainder of this paper is organized as follows. Section "Keywords, distribution of articles, and common technologies in the application of ML techniques in business and finance" outlines the literature retrieval process and presents the statistical findings from the literature analysis, including an analysis of common application trends and ML techniques. Section "Machine learning: a brief introduction" introduces fundamental concepts and terminology related to ML. Sections "Supervised learning" and "Unsupervised learning" explore in-depth common supervised and unsupervised learning techniques, respectively. Section "Applications of machine learning techniques in business and finance" discusses the most recent applications of ML in business and finance. Section "Critical discussions and future research directions" discusses some limitations of ML in this domain and analyzes future research opportunities. Finally, "Conclusions" section concludes.

Gao *et al. Financial Innovation*   (2024) 10:86

Page 4 of 35

Keywords, distribution of articles, and common technologies in the application of ML techniques in business and finance

The primary focus of this review is to explore the advancements in ML in business- and finance-related fields involving ML applications in various market-related issues, including prices, investments, and customer behaviors. This review employs the following strategies to identify existing literature. Initially, we identify relevant journals known for publishing papers that utilize ML techniques to address business and finance problems, such as the UTD-24. Table 2 lists the keywords used in the literature search. During the search process, we input various combinations of ML keywords and business/finance keywords, such as "support vector machine" and "marketing." By cross-referencing the selected journals and keywords and thoroughly examining the citations of highly cited papers, we aimed to achieve a comprehensive and unbiased representation of the current literature.

After identifying journals and keywords, we searched for articles in the Thomson Reuters Web of Science and Elsevier Scopus databases using the same set of keywords. Once the collection phase was complete, the filtering process was initiated. Initially, duplicate articles were excluded to ensure that only unique articles remained for further analysis. Subsequently, we carefully reviewed the full text of each article to eliminate irrelevant or inappropriate items and thus ensure that the final selection comprised relevant and meaningful literature.

Figure 2 illustrates the process of article selection for the review. In the identification phase, we retrieved 154 articles from the search and identified an additional 37 articles through reference checking. During the second phase, duplicates and inappropriate articles were filtered out, resulting in a total of 68 articles eligible for inclusion in this study. Based on the review of these articles, we categorized them into seven different applications: stock market, marketing, e-commerce, energy marketing, cryptocurrency, accounting, and credit risk management, as depicted in Fig. 3 and Tables 3, 4, 5, 6, 7, 8 and 9. Statistical analyses have revealed that ML research in the business and finance domain is predominantly concentrated in the areas of stock market and marketing. The research on e-commerce, cryptocurrency, and energy market applications is nearly equivalent in quantity. Conversely, articles focusing on

Table 2 Keywords

| Type of keywords	Lists of keywords
ML keywords	Data Science, Machine Learning, Supervised Learning, Unsupervised Learning, Support Vector Machine, k-Nearest Neighbor, Random Forest, Decision Tree, Deep Learning, Convolutional Neural Network, Artificial Neural Network, Recurrent Neural Network, Long-Short Term Memory, Kernel based Extreme Learning, Feed-Forward Deep Network, Wavelet Neural Networks
Business/Finance keywords	Economic, Finance, Business, Marketing, Stock, Commerce, Cryptocurrency, Accounting, Credit, Energy
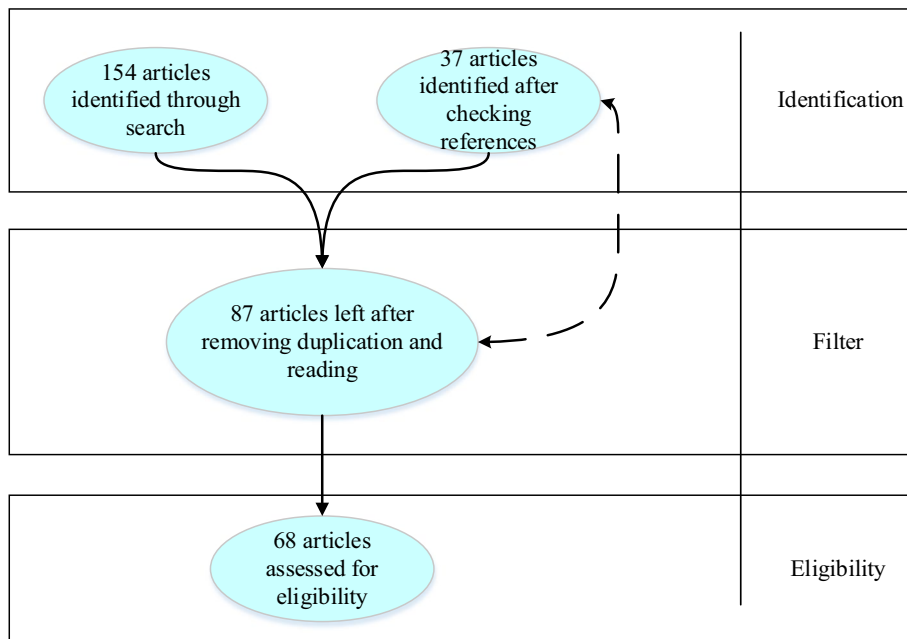
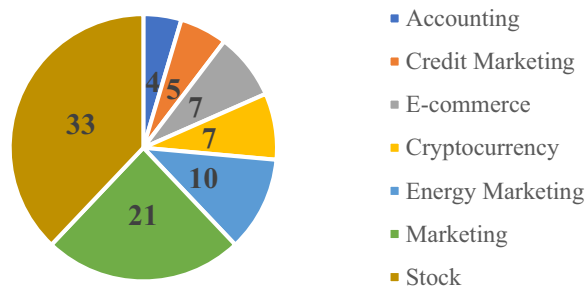**Fig. 2** Flow diagram for article identification and filtering

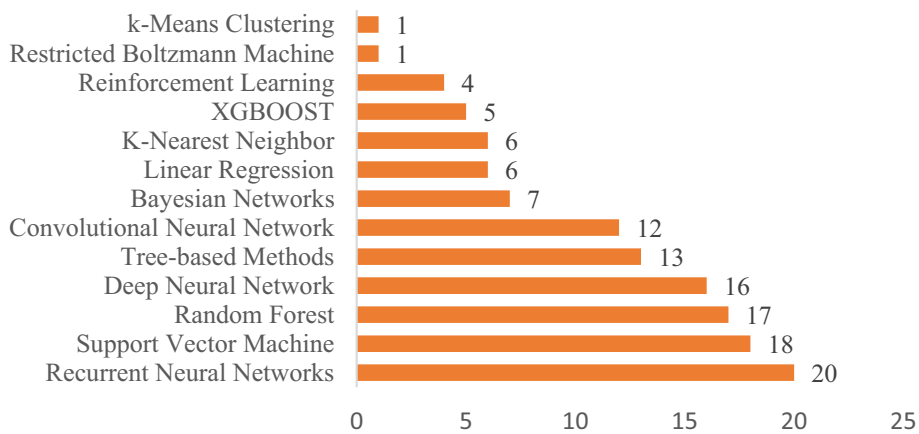**Fig. 3** Number of papers employing ML techniques

Fig. 4 Prominent methods applied in the business and finance domains

Gao *et al. Financial Innovation*   (2024) 10:86

Page 6 of 35

accounting and credit risk management applications are relatively limited. Figure 4 provides a summary of the ML techniques employed in the reviewed articles. Deep learning, support vector machine, and decision tree methods emerged as the most prominent research technologies. In contrast, the application of unsupervised learning techniques, such as k-means and reinforcement learning, were less common.

Machine learning: a brief introduction

This section introduces the basic concepts of ML, including its goals and terminology. Thereafter, we present the model selection method and how to improve the performance.

Goals and terminology

The key objective in various scientific disciplines is to model the relationships between multiple explanatory variables and a set of dependent variables. When a theoretical mathematical model is established, researchers can use it to predict or control desired variables. However, in real-world scenarios, the underlying model is often too complex to be formulated as a closed-form input–output relationship. This complexity has led researchers in the field of ML to focus on developing algorithms (Wu et al. 2008; Chao et al. 2018). The primary goal of these algorithms is to predict certain variables based on other variables or to classify units using limited information; for example, they can be used to classify handwritten digits based on pixel values. ML techniques can automatically construct computational models that capture the intricate relationships present in available data by maximizing the problem-dependent performance criterion or minimizing the error term, which allows them to establish a robust representation of the underlying relationships.

In the context of ML, the sample used to estimate the parameters is usually referred to as a "training sample," and the procedure for estimating the parameters is known as "training." Let N be the sample size, k be the number of features, and q be the number of all possible outcomes. ML can be classified into two main types: supervised and unsupervised. In supervised learning problems, we know both the feature $\mathbf{X}_i = (x_{i1}, ..., x_{ik})$, $i = 1, 2, ..., N$ and the outcome $Y_i = (y_{i1}, y_{i2}, ..., y_{iq})$, where y_{ij} represents the outcome of y_i in the dimension j. For example, in a recommendation system, the quality of product can be scored from 1 to 5, indicating that "q" equals 5. In unsupervised learning problems, we only observe the features $\mathbf{X}_i$ (input data) and aim to group them into clusters based on their similarities or patterns.

Cross-validation, overfitting, and regularization

Cross-validation is frequently used for model selection in ML that is applied to each model; the technique is applied to each model and the one with the lowest expected out-of-sample prediction error is selected.

The ML literature shows significantly higher concern about overfitting than the standard statistics or econometrics literature. In the ML community, the degrees of freedom are not explicitly considered, and many ML methods involve a large number of parameters, which can potentially lead to negative degrees of freedom.

Limiting overfitting is commonly achieved through regularization in ML, which controls the complexity of a model. As stated by Vapnik (2013), the regularization theory

Gao *et al. Financial Innovation*   (2024) 10:86

Page 7 of 35

was one of the first signs of intelligent inference. The complexity of the model describes its ability to approximate various functions. As the complexity increases, the risk of overfitting also increases, whereas less complex and more regularized models may lead to underfitting. Regularization is often implemented by selecting a parsimonious number of variables and using specific functional forms without explicitly controlling for overfitting. Instead of directly optimizing an objective function, a regularization term is added to the objective function, which penalizes the complexity of the model. This approach encourages the model to generalize better and avoids overfitting by promoting simpler and more interpretable solutions.

Here, we provide an example to illustrate how regularization works. The following linear regression model was used:

$$y_{ij} = \sum_{p=1}^{k} b_{pj} x_{ip} + \sigma_j \tag{1}$$

where N is the sample size, k is the numbers of features, and q is the number of all possible outcomes. The variable $y_{ij}(i = 1, 2, ..., N, \quad j = 1, 2, ..., q)$ represents the outcome of y_i in the jth dimension. Additionally, $b_{pj}(p = 1, 2, ..., k, j = 1, 2, ..., q)$ represents the coefficient of feature p in the jth dimension. By using vector notations, $\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_q)^\top$, $\mathbf{b} = (b_{11}, b_{21}, ..., b_{k1}, b_{12}, b_{22}, ..., b_{k2}, ..., b_{1q}, b_{2q}, ..., b_{kq})^\top$ and $Y_i = (y_{i1}, y_{i2}, ..., y_{iq})$, we can rewrite Eq. (1) as follows:

$$Y_i = \mathbf{b}^\top X_i + \boldsymbol{\sigma} \tag{2}$$

where **b** is the solution of

$$\min_{\mathbf{b}} \sum_{i=1}^{N} (Y_i - \mathbf{b}^\top \mathbf{X}_i)^2 + \lambda \left\| \mathbf{b}^\top \right\|_2^2 \tag{3}$$

λ is a penalty parameter that can be selected through out-of-sample cross-validation to optimize the model's out-of-sample predictive performance.

Supervised learning

This section introduces common supervised learning technologies. Compared to traditional statistics, supervised learning methods exhibit certain desired properties when optimizing predictions in large datasets, such as transaction and financial time series data. In business and finance, supervised learning models have proven to be among the most effective tools for detecting credit card fraud (Lebichot et al. 2021). In the following subsections, we briefly describe the commonly used supervised ML methods for business and finance.

Shrinkage methods

The traditional least-squares method often yields complex models with an excessive number of explanatory variables. In particular, when the number of features, *k*, is large compared to the sample size *N*, the least-squares estimator, $\hat{\mathbf{b}}$, does not have good predictive properties, even if the conditional mean of the outcome is linear. To address this problem, regularization is typically used to adjust the estimation parameters dynamically

and reduce the complexity of the model. The shrinkage method is the most common regularization method and can reduce the values of the parameters to be estimated. Shrinkage methods, such as ridge regression (Hoerl and Kennard 1970) and least absolute shrinkage and selection operator (LASSO) (Tibshirani 1996), are linear regression models that add a penalty term to the size of the coefficients. This penalty term pushes the coefficients towards zero, effectively shrinking their values. Shrinkage methods can be effectively used to predict continuous outcomes or classification tasks, particularly when dealing with datasets containing numerous explanatory variables.

Compared to the traditional approach that estimates the regression function using least squares,

$$\hat{\mathbf{b}} = \min_{\mathbf{b}} \sum_{i=1}^{N} (Y_i - \mathbf{b}\mathbf{X}_i)^2 \qquad (4)$$

shrinkage methods add a penalty term that shrinks **b** toward zero, aiming to minimize the following objective function:

$$\min_{\mathbf{b}} \sum_{i=1}^{N} (Y_i - \mathbf{b}\mathbf{X}_i)^2 + \lambda(\|\mathbf{b}\|_q)^{\frac{1}{q}} \qquad (5)$$

where $\|\mathbf{b}\|_q = \sum_{i=1}^{N} |b_i|^q$. In $q = 1$, this formulation leads to a LASSO. However, when $q = 2$ is used, this formulation degenerates ridge regression.

Tree-based method

Regression trees (Breiman et al. 1984) and random forests (Breiman 2001) are effective methods for estimating regression functions with minimal tuning, especially when out-of-sample predictive abilities are required. Considering a sample $(x_{i1}, ..., x_{ik}, Y_i)$ for $i = 1, 2, ..., N$, the idea of a regression tree is to split the sample into subsamples where the regression functions are being estimated. The splits process is sequential and based on feature value x_{ij} exceeding threshold c. Let $R_1(j, c)$ and $R_2(j, c)$ be two sets based on the feature j and threshold c, where $R_1(j, c) = \{\mathbf{X}_i | x_{ij} \leq c\}$ and $R_2(j, c) = \{\mathbf{X}_i | x_{ij} > c\}$. Naturally, the dataset R is divided into two parts, R_1 and R_2, based on the chosen feature and threshold.

Let $c_1 = \frac{1}{|R_1|} \sum_{\mathbf{X}_i \in R_1} x_{ij}$ and $c_2 = \frac{1}{|R_2|} \sum_{\mathbf{X}_i \in R_2} x_{ij}$, where $| \bullet |$ refer to the cardinality of the set. Then we can construct the following optimization model to calculate the errors of the R_1 and R_2 datasets:

$$\min_{j,c}\{\min_{c_1} \sum_{\mathbf{X}_i \in R_1} (Y_i - c_1)^2 + \min_{c_2} \sum_{\mathbf{X}_i \in R_2} (Y_i - c_2)^2\} \qquad (6)$$

For all x_{ij} and threshold $c \in (-\infty, +\infty)$, the method finds the optimal feature j^* and threshold c^* that minimizes errors and splits the sample into subsets based on these criteria. By selecting the best feature and threshold, the method obtains the optimal classification of R_1^* and R_2^*. This process is repeated recursively, leading to further splits that minimize the squared error and improve the overall model performance. However, researchers should be cautious about overfitting, wherein the

Gao *et al. Financial Innovation*   (2024) 10:86

Page 9 of 35

model fits the training data too closely and fails to generalize well to new data. To address this issue, a penalty term can be added to the objective function to encourage simpler and more regularized models. The coefficients of the model are then selected through cross-validation, optimizing the penalty parameter to achieve the best trade-off between model complexity and predictive performance on new, unseen data. This helps prevent overfitting and ensures that the model's performance is robust and reliable.

Random forest builds on the tree algorithm to better estimate the regression function. This approach smooths the regression function by averaging across multiple trees, thus exhibiting two distinct differences. First, instead of using the original sample, each tree is constructed based on a bootstrap sample or a subsample of the data, a technique known as "bagging." Second, at each stage of building a tree, the splits are not optimized over all possible features (covariates) but rather over a random subset of the features. Consequently, feature selection varies in each split, which enhances the diversity of the individual trees.

Deep learning and neural networks

Deep learning and neural networks have been proven to be highly effective in complex settings. However, it is worth noting that the practical implementation of deep learning often demands a considerable amount of tuning compared to other methods, such as decision trees or random forests.

Deep neural networks

As with any other supervised learning methods, deep neural networks (DNNs) can be viewed as a straightforward mapping $y = f(x; \theta)$ from the input feature vector x to the output vector or scalar y, which is governed by the unknown parameters θ. This mapping typically consists of layers that form chain-like structures. Figure 5 illustrates the structure of the DNN. For a DNN with multiple layers, the structure can be represented as

$$y = f^{(k)}(...f^{(2)}(f^{(1)}, \theta_1), \theta_2)..., \theta_k) \tag{7}$$
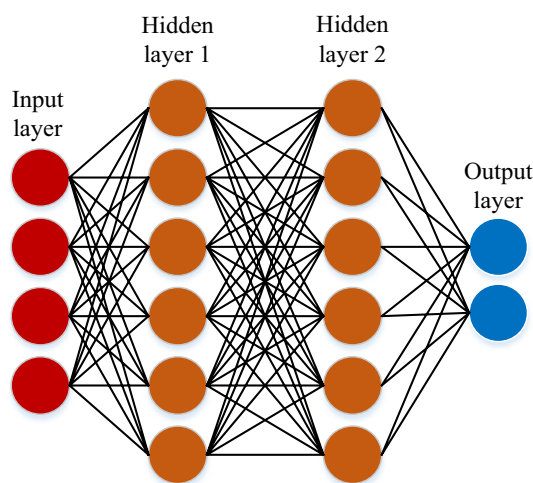

Fig. 5 Structure of DNN

In a fully connected DNN, the ith layer has a structure given by $h^{(i)} = f^{(i)}(x) = g^{(i)}(\mathbf{W}^{(i)}h^{(i-1)} + \mathbf{b}^{(i)})$, where $\mathbf{W}$ is the matrix of unknown parameters and $\mathbf{b}^{(i)}$ is the vector of basis factors. A typical choice for $g^{(i)}$, called the "activation function," can be a rectified linear unit, tanh transformation function, or sigmoid function. The 0th layer $h^{(0)} = x$, which represents the input vector. The row dimension of b or the column dimension of the $\mathbf{W}$ species is the number of neurons in each layer. The weight matrix $\mathbf{W}$ is learned by minimizing a loss function, which can be the mean squared error for regression tasks or the cross-entropy for classification tasks. In particular, when the DNN has one layer, y is scalar. The activation function is set to linear or logistic, and we obtain a linear or logistic regression.

Convolutional neural networks

Although neural networks have many different architectures, the two most classical and relevant are convolutional neural networks (CNNs) and recurrent neural networks (RNNs). A classical CNN structure, which contains three main components—convolutional, pooling, and fully connected layers—is shown in Fig. 6. In contrast to the previously mentioned fully connected structure, in the convolutional layer, each neuron connects with only a small fraction of the neurons from the former layer; however, they share the same parameters. Therefore, sparse connections and parameter sharing significantly reduces the number of estimated parameters.

Different layers play different roles in the training process and are introduced in more detail as follows:

Convolutional layer: This layer comprises a collection of trained filters that are used to extract features from the input data. Assuming that X is the input and there are k filters, the output of the convolutional layer can be formulated as follows:

$$y_j = \sum_i f(x_i * \omega_j + b_j), \quad j = 1, 2, ..., k \tag{8}$$

where ω_j and b_j denote the weights and bias, respectively; f represents the activation function; and $*$ denotes the convolutional operator.

Pooling layer: This layer reduces the features and parameters of the network. The most popular pooling methods are the maximum and average pooling.

CNN are designed to handle one-dimensional time-series data or images. Intuitively, each convolutional layer can be considered a set of filters that move across images or shift along time sequences. For example, some filters may learn to detect textures, whereas others may identify specific shapes. Each filter generates a feature map and the
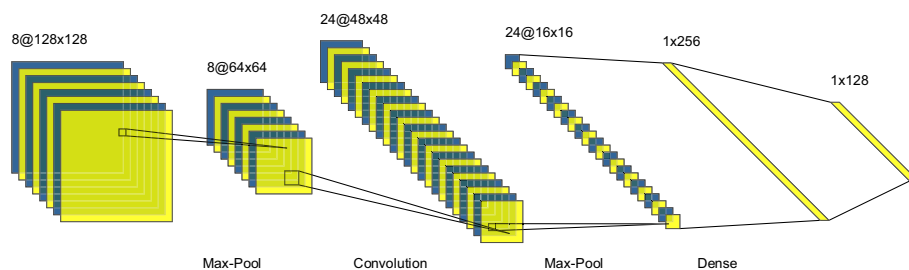

Fig. 6 Structure of CNN

subsequent convolutional layer integrates these features to create a more complex structure, resulting in a map of learned features. Suppose that S is an $p \times p$ window size. Then the average pooling process can be formulated as

$$z = \frac{1}{N} \sum_{(i,j) \in S} x_{ij} \quad i,j = 1, 2, ..., p \tag{9}$$

where x_{ij} is the activation value at location (i, j), and N is the total number of S.

Recurrent neural networks

Recurrent neural networks (RNNs) are well suited for processing sequential data, dynamic relations, and long-term dependencies. RNNs, particularly those employing long short-term memory (LSTM) cells, have become popular and have shown significant potential in natural language processing (Schmidhuber 2015). A key feature of this architecture is its ability to maintain past information over time using a cell-state vector. In each time step, new variables are combined with past information in the cell vector, enabling the RNN to learn how to encode information and determine which encoded information should be retained or forgotten. Similar to CNNs, RNN benefit from parameter sharing, which allows them to detect specific patterns in sequential data.

Figure 7 illustrates the structure of the LSTM network, which contains a memory unit C_t, a hidden state h_t, and three types of gates. Index t refers to the time step. At each step t, the LTSM combines input x_t with the previous hidden state h_{t-1}, calculates the activations of all gates, and updates the memory units and hidden states accordingly.

The computations of LSTM networks are described as follows:

$$
\begin{aligned}
f_t &= \sigma(W_f x_t + \omega_f h_{t-1} + b_f) \\
i_t &= \sigma(W_i x_t + \omega_i h_{t-1} + b_i) \\
O_t &= \sigma(W_o x_t + \omega_o h_{t-1} + b_o) \\
C_t &= f_t \circ C_{t-1} + i_t \circ \sigma_c(W_c x_t + \omega_c h_{t-1} + b_c) \\
h_t &= O_t \circ \tanh(C_t)
\end{aligned}
\tag{10}
$$

where W denotes the weight of the inputs, and ω_f and ω_i represent the weights of the outputs and biases, respectively. The subscript $f, i,$ and O refer to the forget, input, and output gate vectors, respectively. b indicates biases and $\circ$ is an element-wise multiplication.
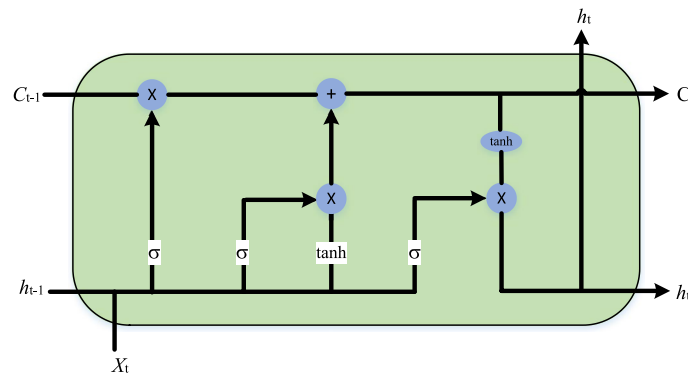
Fig. 7 Structure of LSTM

Wavelet neural networks

Wavelet neural networks (Zhang and Benveniste 1992) use the wavelet function as the activation function, thus combining the advantages of both the wavelet transform and neural networks. The structure of wavelet neural networks is based on backpropagation neural networks, and the transfer function of the hidden layer neuron is the mother wavelet function. For input features $\mathbf{x} = (x_1, ..., x_n)$, the output of the hidden layer can be expressed as follows:

$$h(j) = h_j \left[\frac{\sum_{i=1}^{n} \omega_{ij} x_i - b_j}{a_j} \right], \quad j = 1, 2, ..., m \tag{11}$$

where $h(j)$ is the output value for neuron j, h_j is the mother wavelet function, ω_{ij} is the weight between the input and hidden layers, b_j is the shift factor, and a_j is the stretch factor for h_j.

Support vector machine and kernels

Support vector machines (SVM) are flexible classification methods (Cortes and Vapnik 1995). Let us consider a binary classification problem, where we have an N observation $\mathbf{X}_i$, each with k features, and a binary label $y_i \in \{-1, 1\}$. Subsequently, a hyperplane $x \in \mathbb{R}$ s. t. $w^\top \mathbf{X}_i + b = 0$ is defined, which can be considered a binary classifier $\text{sgn}(w^\top \mathbf{X}_i + b)$. The goal of SVM is to find a hyperplane such that the observations can be separated into two classes: $+1$ and -1. From the hyperplane space, SVM selects the option that maximizes the distance from the closest sample. In an SVM, there is typically a small set of samples with the same maximal distance, which are referred to as "support vectors."

The above-mentioned process can be written as the following optimization model:

$$\min_{\omega, b} \frac{1}{2} \|\omega\|^2 \tag{12}$$

$$s.t. \ Y_i(\omega^\top \mathbf{X}_i + b) \geq 1, \quad \text{for all } i = 1, ..., N \tag{13}$$

To solve the above optimization model, we rewrite it in terms of Lagrangian multipliers as follows:

$$\min_{\alpha, \omega, b} \left\{ \frac{1}{2} \|\omega\|^2 - \sum_{i=1}^{N} \alpha_i (Y_i(\omega^\top \mathbf{X}_i + b) - 1) \right\} \\ s.t. \quad \alpha_i \geq 0 \tag{14}$$

where α_i is the Lagrangian multiplier of the original restriction and $Y_i(\omega^\top \mathbf{X}_i + b) \geq 1$. The model above is equivalent to

$$\max_{\boldsymbol{\alpha}} \left\{ \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j Y_i Y_j \mathbf{X}_i^\top \mathbf{X}_j \right\} \\ s.t. \quad \begin{cases} \alpha_i \geq 0 \\ \sum_{i=1}^{N} \alpha_i Y_i = 0 \end{cases} \tag{15}$$

We can obtain the Lagrangian multiplier $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_N)$ from Model (15), and then $\widehat{b}$ can be solved from $\sum_{i=1}^{N} \hat{\alpha}_i(Y_i(\omega^\top \mathbf{X}_i + b) - 1) = 0$. Furthermore, we can obtain the classifier:

$$f(x) = \text{sgn}(\sum Y_i \hat{\alpha}_i \mathbf{X}_i^\top x + \widehat{b}) \tag{16}$$

Traditional SVM assumes linearly separable training samples. However, SVM can also deal with non-linear cases by mapping the original covariates to a new feature space using the function $\phi(\mathbf{X}_i)$ and then finding the optimal hyperplane in this transformed feature space; that is, $f(x_i) = \omega^\top \phi(x_i) + b$. Thus, the optimization problem in the transformed feature space can be formulated as

$$\max_{\alpha} \left\{ \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j Y_i Y_j K(\mathbf{X}_i, \mathbf{X}_j) \right\}$$
$$s.t. \quad \begin{cases} C \geq \alpha_i \geq 0 \\ \sum_{i=1}^{N} \alpha_i Y_i = 0 \end{cases} \tag{17}$$

where $K(\mathbf{X}_i, \mathbf{X}_j) = \phi(\mathbf{X}_i)^\top \phi(\mathbf{X}_j)$. The kernel function $K(\bullet)$ can be linear, polynomial, or sigmoid. Once the kernel function is determined, we can solve for the value of the Lagrangian multiplier α. Then $\widehat{b}$ can be solved from $\sum_{i=1}^{N} \hat{\alpha}_i(Y_i(\omega^\top \mathbf{X}_i + b) - 1) = 0$, which allows us to derive the classifier:

$$f(x) = \text{sgn}(\sum Y_i \alpha_i K(\mathbf{X}_i, x) + b) \tag{18}$$

Bayesian classifier

A Bayesian network is a graphical model that represents the probabilistic relationships among a set of features (Friedman et al. 1997). The Bayesian network structure S is a directed acyclic graph. Formally, a Bayesian network is a pair $B = \langle G, \Theta \rangle$, where G is a directed acyclic graph whose nodes represent the random variable $(X_1, ..., X_n)$, whose edges represent the dependencies between variables, and Θ is the set of parameters that quantify the graph.

Assuming that there are q labels; that is, $\mathbf{Y} = \{c_1, ..., c_q\}$, λ_{ij} is the loss caused by misclassifying the sample with the true label c_j as c_i, and $\mathbb{X}$ represents the sample space. Then, based on the posterior probability $P(c_i|\mathbf{x})$, we can calculate the expected loss of classifying sample $\mathbf{x}$ into the label c_i as follows:

$$R(c_i|\mathbf{x}) = \sum_{j=1}^{N} \lambda_{ij} P(c_j|\mathbf{x}) \tag{19}$$

Therefore, the aim of the Bayesian classifier is to find a criterion $h : \mathbb{X} \to \mathbf{Y}$ that minimizes the total risk

$$R(h) = E_{\mathbf{x}}[R(h(\mathbf{x})|\mathbf{x})] \tag{20}$$

Obviously, for each sample $\mathbf{x}$, when h can minimize the conditional risk $R(h(\mathbf{x})|\mathbf{x})$, the total risk $R(h)$ will also be minimized. This leads to the concept of Bayes decision rules:

to minimize the total risk, we need to classify each sample into the label that minimizes the conditional risk $R(h(\mathbf{x})|\mathbf{x})$, namely

$$h^*(\mathbf{x}) = \arg \min_{c \in \mathbf{Y}} R(c|\mathbf{x}) \tag{21}$$

We then used h^* as the Bayes-optimal classifier and $R(h^*)$ as the Bayes risk.

K-nearest neighbor

The K-nearest neighbor (KNN) algorithm is a lazy-learning algorithm because it defers to the induction process until classification is required (Wettschereck et al. 1997). The lazy-learning algorithm requires less computation time during the training process compared to eager-learning algorithms such as decision trees, neural networks, and Bayes networks. However, it may require additional time during the classification phase.

The kNN algorithm is based on the assumption that instances close to each other in a feature space are likely to have similar properties. If instances with the same classification label are found nearby, an unlabeled instance can be assigned the same class label as its nearest neighbors. kNN locates the k-nearest instances to the unlabeled instance and determines its label by observing the most frequent class label among these neighbors.

The choice of k significantly affects the performance of the kNN algorithm. Let us discuss the performance of kNN during $k = 1$. Given sample $\mathbf{x}$ and its nearest sample $\mathbf{z}$, the probability of error can be expressed as follows:

$$P(err) = 1 - \sum_{c \in \mathbf{Y}} P(c|\mathbf{x})P(c|\mathbf{z}) \tag{22}$$

Suppose the samples are independent and identically distributed. For any $\mathbf{x}$ and any positive number δ, there always exists at least one sample $\mathbf{z}$ within a distance of δ from $\mathbf{x}$. Let $c^*(\mathbf{x}) \arg \min_{c \in \mathbf{Y}} P(c|\mathbf{x})$ be the outcome the Bayes optimal classifier. Then we have:

$$\begin{aligned}
P(err &= 1 - \sum_{c \in \mathbf{Y}} P(c|\mathbf{x})P(c|\mathbf{z}) \\
&\approx 1 - \sum_{c \in \mathbf{Y}} P^2(c|\mathbf{x}) \\
&\leq 1 - \sum_{c \in \mathbf{Y}} P^2(c^*|\mathbf{x}) \\
&\leq 2(1 - \sum_{c \in \mathbf{Y}} P^2(c^*|\mathbf{x}))
\end{aligned} \tag{23}$$

According to (23), despite the simplicity of kNN, the generalization error is no more than twice that of the Bayes-optimal classifier.

Unsupervised learning

In unsupervised learning, researchers can only access observations without any labeled information, and their primary interest lies in partitioning a sample into subsamples or clusters. Unsupervised learning methods are particularly useful in descriptive tasks because they aim to find relationships in a data structure without measuring the outcomes. Several approaches commonly used in business and finance research fall under

the umbrella of unsupervised learning, including k-means clustering and reinforcement learning. Accordingly, unsupervised learning can be used in qualitative business and finance. For example, it can be particularly beneficial during stakeholder analysis, when stakeholders must be mapped and classified by considering certain predefined attributes. It can also be useful for customer management. A company can employ an unsupervised ML method to cluster guests, which influences its marketing strategy for specific groups and leads to a competitive advantage. This section introduces unsupervised learning technologies that are widely used in business and finance.

K-means clustering

The K-means algorithm aims to find K points in the sample space and classify the samples that are closest to these points. Using an iterative method, the values of each cluster center are updated step-by-step to achieve the best clustering results. When partitioning the feature space into K clusters, the k-means algorithm selects centroids and assigns observations to clusters based on their proximity to them. $b_1, ..., b_k$. The algorithm proceeds as follows. First, we begin with the K centroids $b_1, ..., b_k$, which are initially scattered throughout the feature space. Next, in accordance with the chosen centroids, each observation is assigned to clusters that minimize the distance between the observation and the centroid of the cluster:

$$C_i = \underset{c \in \{1,2,...,k\}}{\arg \min} \left\| \mathbf{X}_i - b_c \right\|^2 \tag{24}$$

Next, we update the centroid by computing the average of X_i across each cluster:

$$b_c = \sum_{i:C_i=c} \mathbf{X}_i \Bigg/ \sum_{i=1}^{k} I(C_i = c) \tag{25}$$

where $I(\bullet)$ is the indicative function. When choosing the number of clusters, K, we must exercise caution because no cross-validation method is available to compare the values.

Reinforcement learning

Reinforcement learning (RL) draws inspiration from the trial-and-error procedure conducted by Thorndike in his 1898 study of cat behavior. Originating from animal learning, RL aims to mimic human behavior by making decisions that maximize profits through interactions with the environment. Mnih et al. (2015) proposed deep RL by employing a deep Q-network to create an agent that outperformed a professional player in a game and further advanced the field of RL.

In deep RL, the learning algorithm plays an essential role in improving efficiency. These algorithms can be categorized into three types: value-based, policy-based, and model-based RL, as illustrated in Fig. 8.

RL consists of four components—agent, state, action and reward—with the agent as its core. When an action leads to a profitable state, it receives a reward, otherwise, it is discouraged. In RL, an agent is defined as any decision-maker, while everything else is considered the environment. The interactions between the environments and the agents are described by state *s*, action *a*, and reward *r*. At time step *t*, the environment is in

state s_t, and the agent takes action a_t. Consequently, the environment transitions to state s_{t+1} and rewards agent r_{t+1}.

The agent's decision is formalized by a policy π, which maps state s to action a. This is deterministic when the probability of choosing action a in state s equals one (i.e., $\pi(a|s) = p(a|s) = 1$). In contrast, it is stochastic when $p(a|s) < 1$ is used. Policy π can be defined as the probability distribution of all actions selected from a certain s, as follows:

$$\begin{aligned} \pi &= \Psi(s) \\ &= \left\{ p(a_i|s) \middle| \forall a_i \in \Delta_\pi \wedge \sum_i p(a_i|s) = 1 \right\} \end{aligned} \tag{26}$$

where Δ_π represents all possible actions of π.

In each step, the agent receives an immediate reward r_{t+1} until it reaches the final state s_T. However, the immediate reward does not ensure a long-term profit. To address this, a generalized return value is used at time step t, defined as R_t:

$$R_t = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + ... + \gamma^{T-t-1} r_T \tag{27}$$

where $0 \le \gamma \le 1$. The agents become more farsighted when γ approaches 1, and more shortsighted when it approaches 0.

The next step is to define a score function V to estimate the goodness of the state:

$$V_\pi(s) = E[R_t | s_t = s, \pi] \tag{28}$$

Then, we determine the goodness of a state-action pair (s, a):

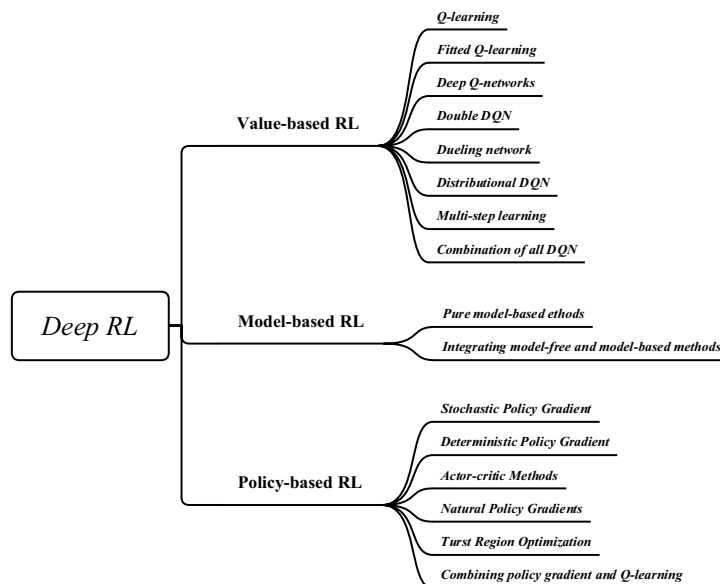$$Q_\pi(s, a) = E[R_t | s_t = s, a_t = a, \pi] \tag{29}$$

Fig. 8 Learning algorithm-based reinforcement learning

Finally, we access the goodness between two policies:

$$\pi \leq \pi' \Leftrightarrow [V_\pi(s) \leq V_{\pi'}(s), \forall s] \vee [Q_\pi(s,a) \leq Q_{\pi'}(s,a), \forall(s,a)] \tag{30}$$

Finally, we can expand $V_\pi(s)$ and $Q_\pi(s,a)$ through R_t to represent the relationship between s and s_{t+1} as

$$V_\pi(s) = \sum_a \pi(s,a) \sum_{s'} p(s'|s,a)(W_{s \to s'|a} + \gamma V_\pi(s')) \tag{31}$$

and

$$Q_\pi(s,a) = \sum_{s'} p(s'|s,a)(W_{s \to s'|a} + \gamma \sum_{a'} \pi(s',a')Q_\pi(s',a')) \tag{32}$$

where $W_{s \to s'|a} = E[r_{t+1}|s_t = s, a_t = a, s_{t+1} = s']$. By solving (31) and (32), we obtain V and S, respectively.

Restricted Boltzmann machines

As Fig. 9 shows, a restricted Boltzmann machine (RBM) can be considered an undirected neural network with two layers, called the "hidden" and "visible" layers. Hidden layers are used to detect the features, whereas visible layers are used to train the input data. Given the n visible layers v and m hidden layers h, the energy function is given by

$$E(v,h) = -\sum_{i=1}^n a_i v_i - \sum_{j=1}^m b_j h_j - \sum_{i=1}^n \sum_{j=1}^m \alpha_{ij} v_i h_j \tag{33}$$

where α_{ij} is the weight between the unit i j, and a_i and b_j are the biases for v and h, respectively.

Applications of machine learning techniques in business and finance

This section considers the application fields in the following categories: marketing, stock market, e-commerce, cryptocurrency, finance, accounting, credit risk management, and energy economy. This study reviews the application status of ML in these fields.

Marketing

ML is an innovative technology that can potentially improve forecasting models and assist in management decision-making. ML applications can be highly beneficial in the marketing domain because they rely heavily on building accurate predictive models from databases. Compared to the traditional statistical approach for forecasting consumer
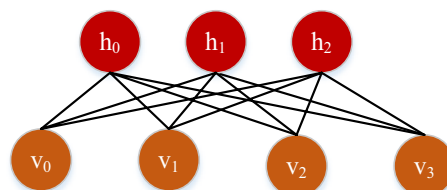

Fig. 9 Structure of RBM

behavior, researchers have recently applied ML technology, which offers several distinctive advantages for data mining with large, noisy databases (Sirignano and Cont 2019). An early example of ML in marketing can be found in the work of Zahavi and Levin (1997), who used neural networks (NNs) to model consumer responses to direct marketing. Compared with the statistical approach, simple forms of NNs are free from the assumptions of normality or complete data, making them particularly robust in handling noisy data. Recently, as shown in Table 3, ML techniques have been predominantly used to study customer behaviors and demands. These applications enable marketers to gain valuable insights and make data-driven decisions to optimize marketing strategies.

Table 3 Machine learning techniques in marketing

References	Research issues	Data source	Machine learning techniques	Journal
Cui et al. (2006)	Customer Behavior	Marketing data set	Bayesian Networks	Management Science
Dingli et al. (2017)	Customer Behavior	A supermarket	RBM, CNN	International Journal of Machine Learning and Computing
Cui et al. (2021)	Customer Demand	Synthetic data	SVM, Random Forest	
Zhang et al. (2020a, b)	Customer Demand	GPS sensor data	Bayesian Networks	Information Systems Research
Timoshenko and Hauser (2019)	Customer Demand	Customer generated data	CNN	Marketing Science
Ladyzynski et al. (2019)	Customer Demand	Time series data of customers	Random Forest, DNN	Expert Systems with Applications
Ferreira et al. (2016)	Customer Demand	Sales transaction data	Regression Tree	Manufacturing & Service Operations Management
Yoganarasimhan(2020)	Customer Demand	Search activity data from Kaggle	Gradient Boosted Decision Trees	Management Science
Padilla and Ascarza (2021)	Customer Relationship	Transactional data of customers	Bayesian Networks	Journal of Marketing Research
Lima et al. (2021)	Customer Demand	ERP system	Linear regression, DNN, SVM, Regression Trees, kNN	INFORMS Journal on Applied Analytics
Boughanmi and Ansari (2021)	Customer Demand	Cross section data of product and user generated information	Bayesian Networks	Journal of Marketing Research
Paolanti et al. (2019)	Production Promotion	Visual and textual retail dataset	CNN	Robotics and Autonomous Systems
Pfeiffer et al.(2020)	Customer Behavior	Field experiment data	SVM, Random Forest	Information Systems Research
Agarwal (2022)	Sentiment Analysis	Social media applications	RNN, CNN	Global Business Review
Zhu et al. (2021)	Customer Demand	Historical sales data	RNN	Production and Operation Management
Melancon et al. (2021)	Customer Demand	Supply chain data	Extreme Gradient Boosting	INFORMS Journal on Applied Analytics
Rafieian and Yoganarasimhan (2021)	Targeting	Mobile in-app advertising network	Extreme Gradient Boosting	Marketing Science
Simester et al. (2020)	Targeting	A large retailer	LASSO, kNN, Kernel Regression, SVM,	Management Science

Consumer behavior refers to the actions taken by consumers to request, use, and dispose of consumer goods, as well as the decision-making process that precedes and determines these actions. In the context of direct marketing, Cui et al. (2006) proposed Bayesian networks that learn by evolutionary programming to model consumer responses to direct marketing using a large direct marketing dataset. In the supply chain domain, Melancon et al. (2021) used gradient-boosted decision trees to predict service-level failures in advance and provide timely alerts to planners for proactive actions. Regarding unsupervised learning in consumer behavior analysis, Dingli et al. (2017) implemented a CNN and an RBM to predict customer churn. However, they found that their performance was comparable to that of supervised learning when introducing added complexity in specific operations and settings. Overall, ML techniques have demonstrated their potential for understanding and predicting consumer behavior, thereby enabling businesses to make informed decisions and optimize their marketing strategies (Machado and Karray 2022; Mao and Chao 2021).

Predicting consumer demand plays a critical role in helping enterprises efficiently arrange production and generate profits. Timoshenko and Hauser (2019) used a CNN to facilitate qualitative analysis by selecting the content for an efficient review. Zhang et al. (2020a, b) used a Bayesian learning model with a rich dataset to analyze the decision-making behavior of taxi drivers in a large Asian city to understand the key factors that drive the supply side of urban mobility markets. Ferreira et al. (2016) employed ML techniques to estimate historical lost sales and predict future demand for new products. For the application of consumer demand-level prediction, most of the research we reviewed used supervised learning technologies because learning consumer consumption preferences requires historical data of consumers, and only clustering consumers is insufficient to predict their consumption levels.

Stock market

ML applications in the stock market have gained immense popularity, with the majority focusing on financial time series for stock price predictions. Table 4 summarizes the reviewed articles that employed ML methods in stock market studies, including references, research objectives, data sources, applied techniques, and journals. Investing in the stock market can be highly profitable but also entails risk. Therefore, investors always try to determine and estimate stock values before taking any action. Researchers have mostly used ML techniques to predict stock prices (Bennett et al. 2022; Moon and Kim 2019). However, predicting stock values can be challenging due to the influence of uncontrollable economic and political factors that make it difficult to identify future market trends. Additionally, financial time-series data are often noisy and nonstationary, rendering traditional forecasting methods less reliable for stock value predictions. Researchers have explored ML in sentiment analysis to identify future trends in the stock market (Baba and Sevil 2021). Furthermore, other studies have focused on objectives such as algorithmic trading, portfolio management, and S&P 500 index trend prediction using ML techniques (Cuomo et al. 2022; Go and Hong 2019).

Various ML techniques have been successfully applied for stock price predictions. Fischer and Krauss (2018) applied LSTM networks to predict the out-of-sample directional movements of the constituent stocks of the S&P 500 from 1992 to 2015, demonstrating

Gao *et al. Financial Innovation*   (2024) 10:86

Page 20 of 35

Table 4 Machine learning techniques in the stock market

References	Research issues	Data source	Machine learning techniques	Journal
Wang et al. (2020)	Portfolio Management	Stock data from the UK Stock Exchange 100 Index	LSTM, SVM, DNN	Expert Systems with Applications
Lei. (2020)	Algorithmic Trading	Day-level real transaction data of U.S. stocks	RL	Expert Systems with Applications
Vo et al. (2019)	Socially Responsible Investments	Financial stock prices and public ESG rating	LSTM	Decision Support Systems
Sabeena and Venkata (2019)	Stock Price Prediction	Online news data	RNN, CNN	International Journal of Engineering and Advanced Technology
Goncalves et al. (2019)	Stock Price Prediction	Betfair betting exchange	CNN, DNNC, LSTM	Information Economics and Policy
Moews et al. (2019)	Stock Price Prediction	Stock market information	DNN	Expert Systems with Applications
Song et al. (2019)	Stock Price Prediction	Stock price data	DNN	Applied Intelligence
Long et al. (2019)	Stock Price Prediction	Financial time series	CNN, RNN	Knowledge-Based Systems
Tashiro et al. (2019)	Stock Price Prediction	Tokyo Stock Exchange	CNN	Quantitative Finance
Sirignano and Cont (2019)	Stock Price Prediction	Financial time series	LSTM	Quantitative Finance
Weng et al. (2018)	Stock Price Prediction	Time series stock market data	SVR, Random Forest	Expert Systems with Applications
Sohangir et al. (2018)	Sentiment Analysis	Social media	CNN, LSTM	Journal of Big Data
Fischer and Krauss (2018)	Stock Price Prediction	S&P 500 index	LSTM, RF, DNN	European Journal of Operational Research
Yan and Ouyang (2018)	Stock Price Prediction	Shanghai Composite Index	LSTM, SVM, kNN	Wireless Personal Communications
Kim et al. (2018)	Stock Price Prediction	Financial time series	Multi-Agent Collaborated Network	International Journal of Grid and Distributed Computing
Chong et al. (2017)	Stock Price Prediction	Stock market in South Korea	DNN	Expert Systems with Applications
Dingli and Fournier (2017)	Stock Price Prediction	Financial time series	CNN	International Journal of Machine Learning and Computing
Singh and Srivastava (2017)	Stock Price Prediction	Financial time series	DNN	Multimedia Tools and Applications
Bao et al. (2017)	Stock Price Prediction	Stock indices	LSTM	PLoS One
Ahmadi et al. (2018)	Stock Price Prediction	Financial time series	SVM	Expert Systems with Applications
Johari et al. (2018)	Stock Price Prediction	Financial time series	SVM	International Journal of Engineering & Technology
Wu et al. (2021)	Hedge Fund Return Prediction	Individual hedge funds	LASSO, Random Forest, Gradient Boosting, DNN	Management Science
Bianchi et al. (2021)	Bond Risk Premiums	Zero-coupon Treasury yield curve data set	Regression Tree, DNN	The Review of Financial Studies

Gao *et al. Financial Innovation*   *(2024) 10:86*

Page 21 of 35

Table 4 (continued)

References	Research issues	Data source	Machine learning techniques	Journal
Gu et al. (2020)	Asset Pricing	S&P 500	Liner Regression, DNN	The Review of Financial Studies

that LSTM networks outperform memory-free classification methods. Wu et al. (2021) applied LASSO, random forest, gradient boosting, and a DNN to cross-sectional return predictions in hedge fund selection and found that ML techniques significantly outperformed four styles of hedge fund research indices in almost all situations. Bao et al. (2017) fed high-level denoising features into the LSTM to forecast the next day's closing price. Sabeena and Venkata (2019) proposed a modified adversarial-network-based framework that integrated a gated recurrent unit and a CNN to acquire data from online financial sites and processed the obtained information using an adversarial network to generate predictions. Song et al. (2019) used deep learning methods to predict future stock prices. Sohangir et al. (2018) applied several NN models to stock market opinions posted on StockTwits to determine whether deep learning models could be adapted to improve the performance of sentiment analysis on StockTwits. Bianchi et al. (2021) showed that extreme trees and NNs provide strong statistical evidence in favor of bond return predictability. Vo et al. (2019) proposed a deep responsible investment portfolio model containing an LSTM network to predict stock returns. All of these stock price applications use supervised learning techniques and financial time-series data to supervise learning. In contrast, it is challenging to apply unsupervised learning methods, particularly clustering, in this domain (Chullamonthon and Tangamchit 2023). However, RL still has certain applications in the stock markets. Lei (2020) combined deep learning and RL models to develop a time-driven, feature-aware joint deep RL model for financial time-series forecasting in algorithmic trading, thus demonstrating the potential of RL in this domain.

Additionally, the evidence suggests that hybrid LSTM methods can outperform other single-supervised ML methods in certain scenarios. Thus, in applying ML to the stock market, researchers have explored the combination of LSTM with different methods to develop hybrid models for improved performance. For instance, Tamura et al. (2018) used LSTM to predict stock prices and reported that the accuracy test results outperformed those of other models, indicating the effectiveness of the hybrid LSTM approach in stock price prediction.

Researchers have explored various hybrid approaches that combine wavelet transforms and LSTM with other techniques to predict stock prices and financial time series. Bao et al. (2017) established a new method for predicting stock prices that integrated wavelet transforms, stacked autoencoders, and LSTM. In the first stage, they eliminate noise to decompose the stock price time series. In the next stage, predictive features for the stock price are created. Finally, LSTM is applied to predict the next day's closing price based on the features of the previous stage. The authors claimed that their model outperformed state-of-the-art models in terms of predictive accuracy and profitability. To address the non-linearity and non-stationary characteristics of financial time series, Yan and Ouyang (2018) integrated wavelet analysis with LSTM to forecast the daily

Gao *et al. Financial Innovation*   (2024) 10:86

Page 22 of 35

closing price of the Shanghai Composite Index. Their proposed model outperformed multiple layer perceptron, SVM, and KNN with respect to finding patterns in financial time-series data. Fang et al. (2019) developed a methodology to predict exchange trade–fund option prices by integrating LSTM with support vector regression (SVR). They used two LSTM-SVR models to model the final transaction price. In the second generation of LSTM-SVR, the hidden state vectors of the LSTM and the seven factors affecting the option price were considered as SVR inputs. Their proposed model outperformed other methods, including LSTM and RF, in predicting option prices.

E-commerce

Online shopping, which allows users to purchase products from companies via the Internet, falls under the umbrella of e-commerce. In today's rapidly evolving online shopping landscape, companies employ effective methods to recognize their buyers' purchasing patterns, thereby enhancing their overall client experience. Customer reviews play a crucial role in this process as they are not only utilized by companies to improve their products and services but also by customers to assess the quality of a product and make informed purchase decisions (Da et al. 2022). Consequently, the decision-making process is significantly improved through analysis of reviews that provide valuable insights to customers.

Traditionally, enterprises' e-commerce strategic planning involves assessing the performance of organizational e-commerce adoption behavior at the strategic level. In this context, the decision-making process exhibits typical behavioral characteristics. With regard to organizations' adoption of technology, it is important to note that the entity adopting the technology is no longer an individual but the organization as a whole. However, technology adoption decisions are still made by people within an organization, and these decisions are influenced by individual cognitive factors (Zha et al. 2021). Individuals involved in the decision-making process have their own perspectives, beliefs, and cognitive biases, which can significantly impact an organization's technology adoption choices and strategies (Li et al. 2019; Xu et al. 2021). Therefore, the behavioral perspective of technology acceptance provides a new perspective for e-commerce strategic planning research. With the development of ML, research on technology acceptance has been hindered by the limitations of traditional strategic e-commerce planning. Different general models of information technology acceptance behaviors are commonly explored.

Table 5 provides a summary of the aforementioned studies. Cui et al. (2021) constructed an e-commerce product marketing model based on an SVM to improve the marketing effects of e-commerce products. Pang and Zhang (2021) built an SVM model to more effectively solve the decision support problem of e-commerce strategic planning. To increase buyers' trust in the quality of the products and encourage online purchases, Saravanan and Charanya (2018) designed an algorithm that categorizes products based on several criteria, including reviews and ratings from other users. They proposed a hybrid feature-extraction method using an SVM to classify and separate products based on their features, best product ratings, and positive reviews. Wang et al. (2018a, b, c) employed LSTM to improve the effectiveness and efficiency of mapping customer requirements to design parameters. The results of their model revealed the superior performance of the RNN over the KNN. Xu et al. (2019) designed an advanced credit risk

Table 5 Machine learning techniques in e-commerce

References	Research issues	Data source	Machine learning techniques	Journal
Saravanan and Charanya (2018)	Product Recommendation	Amazon product dataset	SVM	International Journal of Innovative Technology and Exploring Engineering
Wang et al. (2018a, b, c)	Product Recommendation	Product review data	LSTM	CIRP Annals—Manufacturing Technology
Cui et al. (2021)	Product Marketing	Time series data of customers	SVM, RL	Neural Computing and Applications
Pang and Zhang (2021)	Decision Support System	UCI and StatLog databases	SVM	Information Systems and e-Business Management

evaluation system for e-commerce platforms to minimize the transaction risks associated with buyers and sellers. To this end, they employed a hybrid ML model combined with a decision tree ANN (DT-ANN) and found that it had high accuracy and outperformed other hybrid ML models, such as logistic regression and dynamic Bayesian network. Cai et al. (2018) used deep RL to develop an algorithm to address the allocation of impression problems on e-commerce websites such as www.taobao.com, www.ebay.com, and www.amazon.com. In this algorithm, buyers are allocated to sellers based on their impressions and strategies to maximize the income of the platform. To do so, they applied a gated recurrent unit, and their findings demonstrated that it outperformed a deep deterministic policy gradient. Wu and Yan (2018) claimed that the main assumption of current production recommender models for e-commerce websites is that all historical user data are recorded. In practice, however, many platforms fail to capture such data. Consequently, they devised a list-wise DNN to model the temporal online behavior of users and offered recommendations for anonymous users.

Accounting

In the accounting field, ML techniques are employed to detect fraud and estimate accounting indicators. Most companies' financial statements reflect accounts or disclosure amounts that require estimations. Accounting estimates are pervasive in financial statements and often significantly impact a company's financial position and operational results. The evolution of financial reporting frameworks has led to the increased use of fair value measurements, which necessitates estimation. Most financial statement items are based on subjective managerial estimates and ML has the potential to provide an independent estimate generator (Kou et al. 2021).

Chen and Shi (2020) utilized bagging and boosting ensemble strategies to develop two models: bagged-proportion support vector machines (pSVM) and boosted-pSVMs. Using datasets from LibSVM, they tested their models and demonstrated that ensemble learning strategies significantly enhanced model performance in bankruptcy prediction. Lin et al. (2019) emphasized the importance of finding the best match between

feature selection and classification techniques to improve the prediction performance of bankruptcy prediction models. Their results revealed that using a genetic algorithm as the wrapper-based feature selection method, combined with naïve Bayes and support vector machine classifiers, resulted in remarkable predictive performance. Faris et al. (2019) investigated a combination of resampling (oversampling) techniques and multiple election method features to improve the accuracy of bankruptcy prediction methods. According to their findings, employing the oversampling technique and the AdaBoost ensemble method using a reduced error pruning (REP) tree provided reliable and promising results for bankruptcy prediction.

The earlier studies by Perols (2011) and Perols et al. (2017) were among the first to predict accounting fraud. Two recent studies by Bao et al. (2020) and Bertomeu et al. (2020) used various accounting variables to improve the detection of ongoing irregularities. Bao et al. (2020) employed ensemble learning to develop a fraud-prediction model that demonstrated superior performance compared to the logistic regression and support vector machine models with a financial kernel. Huang et al. (2014) used Bayesian networks to extract textual opinions, and their findings showed that they outperformed dictionary-based approaches, both general and financial. Ding et al. (2020) used insurance companies' data on loss reserve estimates and realizations and documented that the loss estimates generated by ML were superior to the actual managerial estimates reported in financial statements in four out of the five insurance lines examined.

Many companies commission accounting firms to handle accounting and bookkeeping and provide them access to transaction data, documentation, and other relevant information. Mapping daily financial transactions into accounts is one of the most common accounting tasks. Therefore, Jorgensen and Igel (2021) devised ML systems based on random forest to automate the mapping process of financial transfers to the appropriate accounts. Their approach achieved an impressive accuracy of 80.50%, outperforming baseline methods that either excluded transaction text or relied on lexical bag-of-words text representations. The success of ML systems indicates the potential of ML to streamline accounting processes and increase the efficiency of financial transaction' mapping. Table 6 summarizes the ML techniques described in "Accounting" section.

Table 6 Machine learning techniques in accounting

References	Research issues	Data source	Machine learning techniques	Journal
Bao et al. (2020)	Accounting Fraud	Raw accounting numbers	Ensemble Learning	Journal of Accounting Research
Huang et al. (2014)	Investment Behavior	Analyst reports	Bayesian Networks	American Accounting Association
Ding et al. (2020)	Accounting Estimates	Time series insurance payments data	Linear Regression, Random Forest, DNN	Review of Accounting Studies
Jorgensen and Igel (2021)	Financial Transaction	Time series financial transactions data	Random Forest	Intelligent Systems in Accounting, Finance and Management

Table 7 Machine learning techniques in credit risk management

References	Research issues	Data source	Machine learning techniques	Journal
Khandani et al. (2010)	Credit Evaluation	Transactions and credit bureau data	Regression Trees	Journal of Banking & Finance
Ma and Lv (2019)	Credit Risk Prediction	Transactions and credit bureau data	Decision tree	Neural Computing and Applications
Wang et al. (2019)	Credit Risk Prediction	Transactions and credit bureau data	Random Forest, SVM, kNN, K-means	Expert Systems with Applications
Liu et al. (2022)	Default Prediction	Transactions and credit bureau data	kNN, SVM, Random Forest	International Review of Financial Analysis

Credit risk management

The scoring process is an essential part of the credit risk management system used in financial institutions to predict the risk of loan applications because credit scores imply a certain probability of default. Hence, credit scoring modes have been widely developed and investigated for credit approval assessment of new applicants. This process uses a statistical model that considers both the application and performance data of a credit or loan applicant to estimate the likelihood of default, which is the most significant factor used by lenders to prioritize applicants in decision-making. Given the substantial volume of decisions involved in the consumer lending business, it is necessary to rely on models and algorithms rather than on human discretion (Bao et al. 2019; Husmann et al. 2022; Liu et al. 2019). Furthermore, such algorithmic decisions are based on "hard" information, such as consumer credit file characteristics collected by credit bureau agencies.

Supervised and unsupervised ML methods are widely used for credit risk management. Supervised ML techniques are used in credit scoring models to determine the relationships between customer features and credit default risk and subsequently predict classifications. Unsupervised techniques, mainly clustering algorithms, are used as data mining techniques to group samples into clusters (Wang et al. 2019). Hence, unsupervised learning techniques often complement supervised techniques in credit risk management.

Despite the high accuracy of ML, it is not possible to explain its predictions. However, financial institutions must maintain transparency in their decision-making processes. Fortunately, researchers have shown that ML can deduce rules to mitigate a lack of transparency without compromising accuracy (Baesens et al. 2003). Table 7 summarizes the recent applications of ML methods in credit risk management. Liu et al. (2022) use KNN, SVM, and random forest to predict the default probability of online loan borrowers and compare their prediction performance with that of a logistic model. Khandani et al. (2010) applied regression trees to construct non-linear, non-parametric forecasting models for consumer credit risk.

Cryptocurrency

A cryptocurrency is a digital or virtual currency used to securely exchange and transfer assets. Cryptography is used to securely transfer assets, control and regulate the

addition of cryptocurrencies, and secure their transactions (Garcia et al. 2014); hence, the term "cryptocurrency." In contrast to standard currencies, which depend on the central banking system, cryptocurrencies are founded on the principle of decentralized control (Zhao 2021). Owing to its uncontrolled and untraceable nature, the cryptocurrency market has evolved exponentially over a short period. The growing interest in cryptocurrencies in the fields of economics and finance has drawn the attention of researchers in this domain. However, the applications of cryptocurrencies and associated technologies are not limited to financing. There is a significant body of computer science literature that focuses on the supporting technologies of cryptocurrencies, which can lead to innovative and efficient approaches for handling Bitcoin and other cryptocurrencies, as well as addressing their price volatility and other related technologies (Khedr et al. 2021).

Generating an accurate prediction model for such complex problems is challenging. As a result, cryptocurrency price prediction is still in its nascent stages and further research efforts are required to explore this area. In recent years, ML has become one of the most popular approaches for cryptocurrency price prediction owing to its ability to identify general trends and fluctuations. Table 8 presents a survey of cryptocurrency price prediction research using ML methods. Derbentsev et al. (2019) presented a short-term forecasting model to predict the cryptocurrency prices of Ripples, Bitcoin, and Ethereum using an ML approach. Greaves and Au (2015) applied blockchain data to Bitcoin price predictions and employed various ML techniques, including SVM, ANN, and linear and logistic regression. Among the ML classifiers used, the NN classifier with two hidden layers achieved the highest price accuracy of 55%, followed by logistic regression and SVM. Additionally, the research mentioned an analysis using several tree-based models and KNN.

The most recent LSTM networks appear to be more suitable and convenient for handling sequential data, such as time series. Lahmiri and Bekiros (2019) were the first to use LSTM to predict the digital currency prices of the three currencies that were used the most at the time they conducted their study: Bitcoin, Ripple, and digital cash. In their study, long memory was used to assess the market efficiency of cryptocurrencies, and the inherent non-linear dynamics encompassing chaoticity and fractality were examined to gauge the predictability of digital currencies. Chowdhury et al. (2020) applied LSTM

Table 8 Machine learning techniques in cryptocurrency

References	Research issues	Data source	Machine learning techniques	Journal
Lahmiri and Bekiros (2019)	Cryptocurrency forecasting	financial time series	LSTM	Chaos, Solitons and Fractals
Altan et al. (2019)	Cryptocurrency forecasting	financial time series	LSTM	Chaos, Solitons and Fractals
Jiang and Liang (2017)	Cryptocurrency forecasting	financial time series	CNN	
Akyildirim et al. (2021)	Cryptocurrency forecasting	financial time series	SVM, DNN, Random Forest	Annals of Operations Research
Valencia et al. (2019)	Cryptocurrency forecasting	financial time series	DNN, SVM, Random Forest	Entropy
Chowdhury et al. (2020)	Cryptocurrency forecasting	financial time series	Gradient Boosted Trees, DNN, kNN	Physica A: Statistical Mechanics and its Applications

to the indices and constituents of cryptocurrencies to predict prices. Lahmiri and Beki-ros (2019) implemented LSTM to forecast the prices of the three most widely traded cryptocurrencies. Furthermore, Altan et al. (2019) built a novel hybrid forecasting model based on LSTM to predict digital currency time series.

Energy

The existing applications of ML techniques in energy economics can be classified into two major categories: energy price and energy demand prediction. Energy prices typi-cally demonstrate complex features, such as non-linearity, lag dependence, and non-stationarity, which present challenges for the application of simple traditional models (Chen et al. 2018). Owing to their high flexibility, ML techniques can provide superior prediction performance. In energy demand predictions, lagged values of consumption and socioeconomic and technological variables, such as GDP per capita, population, and technology trends, are typically utilized. Table 9 presents a summary of these studies. A critical distinction between "price" and "consumption" prediction is that the latter is not subject to market efficiency dynamics. The prediction of consumption has little effect on the actual consumption of the agents. However, price prediction tends to offset itself by creating opportunities for traders to use this information.

Predicting prices in energy markets is a complicated process because prices are sub-ject to physical constraints on electricity generation and transmission and market power potential (Young et al. 2014). Predicting prices using ML techniques is one of the old-est applications in energy economics. In the early 2000s, a wave of studies attempted to forecast electricity prices using conventional ANN techniques. Ding (2018) combined ensemble empirical mode decomposition and an artificial NN to forecast international crude oil prices. Zhang et al. (2020a, b) employed the LSTM method to forecast day-ahead electricity prices in a deregulated electricity market. They also investigated the intricate dependence structure within the price-forecasting model. Peng et al. (2018)

Table 9 Machine learning techniques in energy marketing

References	Research issues	Data source	Machine learning techniques	Journal
Zhang et al. (2020a, b)	Energy Prices Predict-ing	Electricity market	LSTM	the Institution of Engineering and Technology
Ding (2018)	Energy Prices Predict-ing	Crude oil price series	DNN	Energy
Peng et al. (2018)	Energy Prices Predict-ing	Electricity market	LSTM	Energy
Lago et al. (2018)	Energy Prices Predict-ing	Electricity market	DNN	Applied Energy
Huang and Wang (2018)	Energy Prices Predict-ing	Crude oil price series	Wavelet Neural Network	Energy
Chen et al. (2018)	Energy Demand predicting	Electric load data series	Random Forest, DNN	Energy
Wang et al. (2018a, b, c)	Energy Demand predicting	Energy consumption and energy consump-tion structure data	SVM	Energy
Alobaidi et al. (2018)	Energy Demand predicting	Electric load data series	Ensemble Learning	Applied Energy

applied LSTM with a differential evolution algorithm to predict electricity prices. Lago et al. (2018) first proposed a DNN to improve the predictive accuracy in a local market and then proposed a second model that simultaneously predicts prices from two markets to further improve the forecasting accuracy. Huang and Wang (2018) proposed a model that combines wavelet NNs with random time-effective functions to improve the prediction accuracy of crude oil price fluctuations.

Understanding the future energy demand and consumption is essential for short- and long-term planning. A wide range of users, including government agencies, local development authorities, financial institutions, and trading institutions, are interested in obtaining realistic forecasts of future consumption portfolios (Lei et al. 2020). For demand prediction, Chen et al. (2018) used ridge regression to combine extreme gradient boosting forest and feedforward deep networks to predict the annual household electricity consumption. Wang et al. (2018a, b, c) first built a model using a self-adaptive multi-verse optimizer to optimize the SVM and then employed it to predict China's primary energy consumption.

Critical discussions and future research directions

ML techniques have proven valuable in establishing computational models that capture complex relationships with the available data. Consequently, ML has become a useful tool in business and finance. This section critically discusses the existing research and outlines future directions.

Critical discussions

Although ML techniques are widely employed in business and finance, several issues need to be addressed.

1. Linguistic information is abundant in business and finance, encompassing online commodity comments and investors' emotional responses in the stock market. Nonetheless, the existing research has predominantly concentrated on processing numerical data. When juxtaposed with numerical information, linguistic data harbor intricate characteristics, notably personalized individual semantics (Li et al. 2022a, b; Zhang et al. 2021a, b; Hoang and Wiegratz 2022).

2. The integration of ML into business and finance can lead to interpretability issues. In ML, an interpretable model refers to one in which a human observer can readily comprehend how the model transforms an observation into a prediction (Freitas 2014). Typically, decision-makers are hesitant to accept recommendations generated by ML techniques unless they can grasp the reasoning behind them. Unfortunately, the existing research in business and finance, particularly those employing DNNs, has seldom emphasized the interpretability of their models.

3. Social networks are prevalent in the marketing domain within businesses (Zha et al. 2020). For instance, social networks exist among consumers, whose purchasing behavior is influenced by the opinions of trusted peers or friends. However, the existing research that applies ML to marketing has predominantly concentrated on personal customer attributes, such as personality, purchasing power, and preferences (Dong et al. 2021). Regrettably, the potential impact of social networks and their influence on customer behavior have been largely overlooked in these studies.

4. ML techniques typically focus on exploring the statistical relationships between dependent and independent variables and emphasize feature correlations. However, in the context of business and finance applications, causal relationships exist between variables. For instance, consider a study suggesting that girls who have breakfast tend to have lower weights than those who do not, based on which one might conclude that having breakfast aids in weight loss. However, in reality, these two events may only exhibit a correlation rather than causation (Yao et al. 2021). Causality plays a significant role in ML techniques' performance. However, many current business and finance applications have failed to account for this crucial factor. Ignoring causality may lead to misleading conclusions and hinder accurate modeling of real-world scenarios. Therefore, incorporating causality into ML methodologies within the business and finance domains is essential for enhancing the reliability and validity of predictive models and decision-making processes.

5. In the emerging cryptocurrency field, although traditional statistical methods are simple to implement and interpret, they require many unrealistic statistical assumptions, making ML the best technology in this field. Although many ML techniques exist, challenges remain in accurately predicting cryptocurrency prices. However, most ML techniques require further investigation.

6. In recent years, rapid growth in digital payments has led to significant shifts in fraud and financial crimes (Canhoto 2021; Prusti et al. 2022; Wang et al. 2023). While some studies have shown the effective use of ML in detecting financial crimes, there remains a limitation in the research dedicated to this area. As highlighted by Pourhabibi et al. (2020), the complex nature of financial crime detection applications poses challenges in terms of deploying and achieving the desired detection performance levels. These challenges are manifested in two primary aspects. First, ML solutions encounter substantial pressure to deliver real-time responses owing to the constraints of processing data in real time. Second, in addition to inherent data noise, criminals often attempt to introduce deceptive data to obfuscate illicit activities (Pitropakis et al. 2019). Regrettably, few studies have investigated the robustness and performance of the underlying algorithmic solutions when confronted with data quality issues.

7. In the finance domain, an important limitation of the current literature on energy and ML is that most works highlight the computer science perspective to optimize computational parameters (e.g., the accuracy rate), while finance intuition may be ignored.

Future research directions

Thus, we propose that future research on this topic follow the directions below:

1. As analyzed above, there is abundant linguistic information exists in business and finance. Consequently, leveraging natural language processing technology to handle and analyze linguistic data in these domains represents a highly promising research direction.

2. The amalgamation of theoretical models using ML techniques is an important research topic. The incorporation of interpretable models can effectively reveal the black-box nature of ML-driven analyses, thereby elucidating the underlying reasoning behind the results. Consequently, the introduction of interpretable models into business and finance while applying ML can yield substantial benefits.

3. The interactions and behaviors are often intertwined within social networks, making it crucial to incorporate social network dynamics when modeling their influence on consumer behavior. Introducing the social network aspect into ML models has tremendous potential for enhancing marketing strategies and outcomes (Trandafili and Biba 2013).

4. Causality has garnered increasing attention in the field of ML in recent years. Accordingly, we believe it is an intriguing avenue to explore when applying ML to address problems in business and finance.

5. Further studies need to include all relevant factors affecting market mood and track them over a longer period to understand the anomalous behavior of cryptocurrencies and their prices. We recommend that researchers analyze the use of LSTM models in future research, such as CNN LSTM and encoder–decoder LSTM, and compare the results to obtain future insights and improve price prediction results. In addition, researchers can apply sentiment analysis to collect social signals, which can be further enhanced by improving the quality of content and using more content sources. Another area of opportunity is the use of more specialized models with different types of approaches, such as LSTM networks.

6. Graph NNs and emerging adaptive solutions provide important opportunities for shaping the future of fraud and financial crime detection owing to their parallel structures. Because of the complexity of digital transaction processing and the ever-changing nature of fraud, robustness should be treated as the primary design goal when applying ML to detect financial crimes. Finally, focusing on real-time responses and data noise issues is necessary to improve the performance of current ML solutions for financial crime detection.

7. Currently, the application of unsupervised learning methods in different areas, such as marketing and risk management, is limited. Some problems related to marketing and customer management could be analyzed using clustering techniques, such as K-means, to segment clients by different demographic or behavioral characteristics and by their likelihood of default or switching companies. In energy risk management, extreme events can be identified as outliers using principal component analysis or ranking algorithms.

Conclusions

Having already made notable contributions to business and finance, ML techniques for addressing issues in these domains are significantly increasing. This review discusses advancements in ML in business and finance by examining seven research directions of ML techniques: cryptocurrency, marketing, e-commerce, energy marketing, stock market, accounting, and credit risk management. Deep learning models, such as DNN, CNN, RNN, random forests, and SVM are highlighted in almost every domain of business and finance. Finally, we analyze some limitations of existing studies and suggest

several avenues for future research. This review is helpful for researchers in understanding the progress of ML applications in business and finance, thereby promoting further developments in these fields.

Abbreviations
ML   Machine learning
DNN   Deep neural networks
CNN   Convolutional neural networks
RNN   Recurrent neural networks
LSTM   Long short-term memory
SVM   Support vector machine
kNN   K-nearest neighbor
RL   Reinforcement learning
RBM   Restricted Boltzmann machine
LASSO   Least absolute shrinkage and selection operator

Author contributions
HG, GK and YD contributed to the completion of the idea and writing of this paper. HG, GK and YD contributed to the discussion of the content of the organization and HL and HZ contributed to the improvement of the text of the manuscript. HG and HL contributed to Methodology. XC, and CL contributed to the literature collection of this paper. All authors read and approved the final manuscript.

Availability of data and materials
Not applicable.

Declarations

Competing interests
The authors declare that they have no competing interests.

References
Agarwal S (2022) Deep learning-based sentiment analysis: establishing customer dimension as the lifeblood of business management. Glob Bus Rev 23(1):119–136
Ahmadi E, Jasemi M, Monplaisir L, Nabavi MA, Mahmoodi A, Jam PA (2018) New efficient hybrid candlestick technical analysis model for stock market timing on the basis of the support vector machine and heuristic algorithms of imperialist competition and genetic. Expert Syst Appl 94:21–31
Akyildirim E, Goncu A, Sensoy A (2021) Prediction of cryptocurrency returns using machine learning. Ann Oper Res 297(1–2):34
Alobaidi MH, Chebana F, Meguid MA (2018) Robust ensemble learning framework for day-ahead forecasting of household-based energy consumption. Appl Energy 212:997–1012
Altan A, Karasu S, Bekiros S (2019) Digital currency forecasting with chaotic meta-heuristic bio-inspired signal processing techniques. Chaos Solitons Fractals 126:325–336
Athey S, Imbens GW (2019) Machine learning methods that economists should know about. Annu Rev Econ 11:685–725
Baba B, Sevil G (2021) Bayesian analysis of time-varying interactions between stock returns and foreign equity flows. Financ Innov 7(1):51
Baesens B, Setiono R, Mues C, Vanthienen J (2003) Using Neural Network Rule Extraction and Decision Tables for Credit-Risk Evaluation. Manage Sci 49(3):312–329
Bajari P, Nekipelov D, Ryan SP, Yang MY (2015) Machine learning methods for demand estimation. Am Econ Rev 105(5):481–485
Bao W, Yue J, Rao YL (2017) A deep learning framework for financial time series using stacked autoencoders and long-short term memory. PLoS ONE 12(7):24
Bao W, Lianju N, Yue K (2019) Integration of unsupervised and supervised machine learning algorithms for credit risk assessment. Expert Syst Appl 128:301–315

Bao Y, Ke BIN, Li BIN, Yu YJ, Zhang JIE (2020) Detecting accounting fraud in publicly traded U.S. firms using a machine learning approach. J Acc Res 58(1):199–235

Bennett S, Cucuringu M, Reinert G (2022) Lead–lag detection and network clustering for multivariate time series with an application to the US equity market. Mach Learn 111(12):4497–4538

Bianchi D, Buchner M, Tamoni A (2021) Bond risk premiums with machine learning. Rev Financ Stud 34(2):1046–1089

Boughanmi K, Ansari A (2021) Dynamics of musical success: a machine learning approach for multimedia data fusion. J Mark Res 58(6):1034–1057

Breiman L (2001) Random forests. Mach Learn 45(1):5–32

Breiman L, Friedman J, Olshen R, Stone C (1984) Classification and regression trees. Chapman and Hall, Wadsworth

Cai Q, Filos-Ratsikas A, Tang P, Zhang Y (2018) Reinforcement mechanism design for e-commerce. In: Proceedings of the 2018 world wide web conference, pp 1339–1348

Canhoto AI (2021) Leveraging machine learning in the global fight against money laundering and terrorism financing: an affordances perspective. J Bus Res 131:441–452

Chen KL, Jiang JC, Zheng FD, Chen KJ (2018) A novel data-driven approach for residential electricity consumption prediction based on ensemble learning. Energy 150:49–60

Chao X, Kou G, Li T, Peng Y (2018) Jie Ke versus AlphaGo: a ranking approach using decision making method for large-scale data with incomplete information. Eur J Oper Res 265(1):239–247

Chen Z, Chen W, Shi Y (2020) Ensemble learning with label proportions for bankruptcy prediction. Expert Syst Appl 146:113155

Chen H, Fang X, Fang H (2022) Multi-task prediction method of business process based on BERT and transfer learning. Knowl Based Syst 254:109603

Chen MR, Dautais Y, Huang LG, Ge JD (2017) Data driven credit risk management process: a machine learning approach. Paper presented at the international conference on software and system process Paris, France

Chong E, Han C, Park FC (2017) Deep learning networks for stock market analysis and prediction: methodology, data representations, and case studies. Expert Syst Appl 83:187–205

Chowdhury R, Rahman MA, Rahman MS, Mahdy MRC (2020) An approach to predict and forecast the price of constituents and index of cryptocurrency using machine learning. Physica A 551:17

Chullamonthon P, Tangamchit P (2023) Ensemble of supervised and unsupervised deep neural networks for stock price manipulation detection. Expert Syst Appl 220:119698

Coble KH, Mishra AK, Ferrell S, Griffin T (2018) Big data in agriculture: a challenge for the future. Appl Econ Perspect Policy 40(1):79–96

Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297

Cui G, Wong ML, Lui HK (2006) Machine learning for direct marketing response models: Bayesian networks with evolutionary programming. Manag Sci 52(4):597–612

Cui F, Hu HH, Xie Y (2021) An intelligent optimization method of e-commerce product marketing. Neural Comput Appl 33(9):4097–4110

Cuomo S, Gatta F, Giampaolo F, Iorio C, Piccialli F (2022) An unsupervised learning framework for marketneutral portfolio. Expert Syst Appl 192:116308

Da F, Kou G, Peng Y (2022) Deep learning based dual encoder retrieval model for citation recommendation. Technol Forecast Soc 177:121545

Dastile X, Celik T, Potsane M (2020) Statistical and machine learning models in credit scoring: A systematic literature survey. Appl Soft Comput 91:21

Derbentsev V, Datsenko N, Stepanenko O, Bezkorovainyi V (2019) Forecasting cryptocurrency prices time series using machine learning approach. In: SHS web of conferences, vol 65, p 02001

Ding YS (2018) A novel decompose-ensemble methodology with AIC-ANN approach for crude oil forecasting. Energy 154:328–336

Ding KX, Lev B, Peng X, Sun T, Vasarhelyi MA (2020) Machine learning improves accounting estimates: evidence from insurance payments. Rev Acc Stud 25(3):1098–1134

Dingli A, Fournier KS (2017) Financial time series forecasting - a deep learning approach. Int J Mach Learn Comput 7(5):118–122

Dingli A, Marmara V, Fournier NS (2017) Comparison of deep learning algorithms to predict customer churn within a local retail industry. Int J Mach Learn Comput 7(5):128–132

Dong YC, Li Y, He Y, Chen X (2021) Preference-approval structures in group decision making: axiomatic distance and aggregation. Decis Anal 18(4):273–295

Einav L, Levin J (2014) Economics in the age of big data. Science 346(6210):715–＋

Fang Y, Chen J, Xue Z (2019) Research on quantitative investment strategies based on deep learning. Algorithms 12(2):35

Faris H, Abukhurma R, Almanaseer W, Saadeh M, Mora AM, Castillo PA, Aljarah I (2019) Improving financial bankruptcy prediction in a highly imbalanced class distribution using oversampling and ensemble learning: A case from the Spanish market. Prog Artif Intell 9:1–23

Ferreira KJ, Lee BHA, Simchi-Levi D (2016) Analytics for an online retailer: demand forecasting and price optimization. Manuf Serv Oper Manag 18(1):69–88

Fischer T, Krauss C (2018) Deep learning with long short-term memory networks for financial market predictions. Eur J Oper Res 270(2):654–669

Freitas AA (2014) Comprehensible classification models: a position paper. SIGKDD Explor Newsl 15(1):1–10

Friedman N, Geiger D, Goldszmidt M (1997) Bayesian network classifiers. Mach Learn 29(2–3):131–163

Garcia D, Tessone CJ, Mavrodiev P, Perony N (2014) The digital traces of bubbles: feedback cycles between socio-economic signals in the bitcoin economy. J R Soc Interface 11(99):20140623

Ghoddusi H, Creamer GG, Rafizadeh N (2019) Machine learning in energy economics and finance: a review. Energy Econ 81:709–727

Go YH, Hong JK (2019) Prediction of stock value using pattern matching algorithm based on deep learning. Int J Recent Technol Eng 8:31–35

Gogas P, Papadimitriou T (2021) Machine learning in economics and finance. Comput Econ 57(1):1–4

Goncalves R, Ribeiro VM, Pereira FL, Rocha AP (2019) Deep learning in exchange markets. Inf Econ Policy 47:38–51

Greaves A, Au B (2015) Using the bitcoin transaction graph to predict the price of bitcoin. No Data

Grimmer J (2015) We are all social scientists now: how big data, machine learning, and causal inference work together. PS Polit Sci Polit 48(1):80–83

Gu SH, Kelly B, Xiu DC (2020) Empirical Asset Pricing via Machine Learning. Rev Financ Stud 33(5):2223–2273

Hoang D, Wiegratz K (2022) Machine learning methods in finance: Recent applications and prospects. Eur Financ Manag 29(5):1657–1701

Hoerl AE, Kennard RW (1970) Ridge regression—biased estimation for nonorthogonal problems. Technometrics 12(1):55–000

Huang LL, Wang J (2018) Global crude oil price prediction and synchronization-based accuracy evaluation using random wavelet neural network. Energy 151:875–888

Huang AH, Zang AY, Zheng R (2014) Evidence on the information content of text in analyst reports. Account Rev 89(6):2151–2180

Husmann S, Shivarova A, Steinert R (2022) Company classification using machine learning. Expert Syst Appl 195:116598

Jiang ZY, Liang JJ (2017) Cryptocurrency portfolio management with deep reinforcement learning. In: Paper presented at the intelligent systems conference, London, England

Johari SN, Farid FH, Nasrudin N, Bistamam NL, Shuhaili NS (2018) Predicting Stock Market Index Using Hybrid Intelligence Model. Int J Eng Technol 7:36

Jorgensen RK, Igel C (2021) Machine learning for financial transaction classification across companies using character-level word embeddings of text fields. Intell Syst Account Financ Manag 28(3):159–172

Kamilaris A, Prenafeta-Boldu FX (2018) Deep learning in agriculture: a survey. Comput Electron Agric 147:70–90

Khandani AE, Kim AJ, Lo AW (2010) Consumer credit-risk models via machine-learning algorithms. J Bank Financ 34(11):2767–2787

Khedr AM, Arif I, Raj PVP, El-Bannany M, Alhashmi SM, Sreedharan M (2021) Cryptocurrency price prediction using traditional statistical and machine-learning techniques: a survey. Intell Syst Account Financ Manag 28(1):3–34

Kim JJ, Cha SH, Cho KH, Ryu M (2018) Deep reinforcement learning based multi-agent collaborated network for distributed stock trading. Int J Grid Distrib Comput 11(2):11–20

Kou G, Chao XR, Peng Y, Alsaadi FE, Herrera-Viedma E (2019) Machine learning methods for systemic risk analysis in financial sectors. Technol Econ Dev Eco 25(5):716–742

Kou G, Xu Y, Peng Y, Shen F, Chen Y, Chang K, Kou S (2021) Bankruptcy prediction for SMEs using transactional data and two-stage multiobjective feature selection. Decis Support Syst 140:113429

Ladyzynski P, Zbikowski K, Gawrysiak P (2019) Direct marketing campaigns in retail banking with the use of deep learning and random forests. Expert Syst Appl 134:28–35

Lago J, De Ridder F, Vrancx P, De Schutter B (2018) Forecasting day-ahead electricity prices in Europe: the importance of considering market integration. Appl Energy 211:890–903

Lahmiri S, Bekiros S (2019) Cryptocurrency forecasting with deep learning chaotic neural networks. Chaos Solitons Fractals 118:35–40

Lebichot B, Paldino GM, Siblini W, Guelton LH, Oblé F, Bontempi G (2021) Incremental learning strategies for credit cards fraud detection. Int J Data Sci Anal 12:165–174

Lei ZZ (2020) Research and analysis of deep learning algorithms for investment decision support model in electronic commerce. Electron Commer Res 20(2):275–295

Lei K, Zhang B, Li Y, Yang M, Shen Y (2020) Time-driven feature-aware jointly deep reinforcement learning for financial signal representation and algorithmic trading. Expert Syst Appl 140:14

Li CC, Dong YC, Xu YJ, Chiclana F, Herrera-Viedma E, Herrera F (2019) An overview on managing additive consistency of reciprocal preference relations for consistency-driven decision making and Fusion: Taxonomy and future directions. Inf Fusion 52:143–156

Li CC, Dong YC, Liang H, Pedrycz W, Herrera F (2022a) Data-driven method to learning personalized individual semantics to support linguistic multi-attribute decision making. Omega 111:102642

Li CC, Dong YC, Pedrycz W, Herrera F (2022b) Integrating continual personalized individual semantics learning in consensus reaching in linguistic group decision making. IEEE Trans Syst Man Cybern Syst 52(3):1525–1536

Lima MSM, Eryarsoy E, Delen D (2021) Predicting and explaining pig iron production on charcoal blast furnaces: a machine learning approach. INFORMS J Appl Anal 51(3):213–235

Lin WY, Hu YH, Tsai CF (2012) Machine learning in financial crisis prediction: a survey. IEEE Trans Syst Man Cybern Syst C 42(4):421–436

Lin WC, Lu YH, Tsai CF (2019) Feature selection in single and ensemble learning-based bankruptcy prediction models. Expert Syst 36:e12335

Liu YT, Zhang HJ, Wu YZ, Dong YC (2019) Ranking range based approach to MADM under incomplete context and its application in venture investment evaluation. Technol Econ Dev Eco 25(5):877–899

Liu Y, Yang ML, Wang YD, Li YS, Xiong TC, Li AZ (2022) Applying machine learning algorithms to predict default probability in the online credit market: evidence from China. Int Rev Financ Anal 79:14

Long W, Lu ZC, Cui LX (2019) Deep learning-based feature engineering for stock price movement prediction. Knowl Based Syst 164:163–173

Ma XM, Lv SL (2019) Financial credit risk prediction in internet finance driven by machine learning. Neural Comput Appl 31(12):8359–8367

Machado MR, Karray S (2022) Applying hybrid machine learning algorithms to assess customer risk-adjusted revenue in the financial industry. Electron Commer Res Appl 56:101202

Mao ST, Chao XL (2021) Dynamic joint assortment and pricing optimization with demand learning. Manuf Serv Oper Manag 23(2):525–545

Melancon GG, Grangier P, Prescott-Gagnon E, Sabourin E, Rousseau LM (2021) A machine learning-based system for predicting service-level failures in supply chains. INFORMS J Appl Anal 51(3):200–212

Meng TL, Khushi M (2019) Reinforcement learning in financial markets. Data 4(3):110

Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, Graves A, Riedmiller M, Fidjeland AK, Ostrovski G, Petersen S, Hassabis D (2015) Human-level control through deep reinforcement learning. Nature 518(7540):529–533

Moews B, Herrmann JM, Ibikunle G (2019) Lagged correlation-based deep learning for directional trend change prediction in financial time series. Expert Syst Appl 120:197–206

Moon KS, Kim H (2019) Performance of deep learning in prediction of stock market volatility. Econ Comput Econ Cybern Stud 53(2):77–92

Nanduri J, Jia YT, Oka A, Beaver J, Liu YW (2020) Microsoft uses machine learning and optimization to reduce e-commerce fraud. Informs J Appl Anal 50(1):64–79

Nazareth N, Ramana RYV (2023) Financial applications of machine learning: a literature review. Expert Syst Appl 219:119640

Nguyen TT, Nguyen ND, Nahavandi S (2020) Deep reinforcement learning for multiagent systems: a review of challenges, solutions, and applications. IEEE Trans Cybern 50(9):3826–3839

Nosratabadi S, Mosavi A, Duan P, Ghamisi P, Filip F, Band SS, Reuter U, Gama J, Gandomi AH (2020) Data science in economics: comprehensive review of advanced machine learning and deep learning methods. Mathematics 8(10):1799

Nti IK, Adekoya AF, Weyori BA (2020) A systematic review of fundamental and technical analysis of stock market predictions. Artif Intell Rev 53(4):3007–3057

Ozbayoglu AM, Gudelek MU, Sezer OB (2020) Deep learning for financial applications: a survey. Appl Soft Comput 93:106384

Padilla N, Ascarza E (2021) Overcoming the cold start problem of customer relationship management using a probabilistic machine learning approach. J Mark Res 58(5):981–1006

Pang H, Zhang WK (2021) Decision support model of e-commerce strategic planning enhanced by machine learning. Inf Syst E-Bus Manag 21(1):11

Paolanti M, Romeo L, Martini M, Mancini A, Frontoni E, Zingaretti P (2019) Robotic retail surveying by deep learning visual and textual data. Robot Auton Syst 118:179–188

Peng L, Liu S, Liu R, Wang L (2018) Effective long short-term memory with differential evolution algorithm for electricity price prediction. Energy 162:1301–1314

Perols J (2011) Financial statement fraud detection: An analysis of statistical and machine learning algorithms. Auditing J Pract Th 30:19–50

Perols JL, Bowen RM, Zimmermann C, Samba B (2017) Finding needles in a haystack: using data analytics to improve fraud prediction. Acc Rev 92(2):221–245

Pfeiffer J, Pfeiffer T, Meissner M, Weiss E (2020) Eye-tracking-based classification of information search behavior using machine learning: evidence from experiments in physical shops and virtual reality shopping environments. Inf Syst Res 31(3):675–691

Pitropakis N, Panaousis E, Giannetsos T, Anastasiadis E, Loukas G (2019) A taxonomy and survey of attacks against machine learning. Comput Sci Rev 34:100199

Pourhabibi T, Ong KL, Kam BH, Boo YL (2020) Fraud detection: a systematic literature review of graph-based anomaly detection approaches. Decis Support Syst 133:113303

Prusti D, Behera RK, Rath SK (2022) Hybridizing graph-based Gaussian mixture model with machine learning for classification of fraudulent transactions. Comput Intell 38(6):2134–2160

Rafieian O, Yoganarasimhan H (2021) Targeting and privacy in mobile advertising. Mark Sci 40(2):193–218

Raj MP, Swaminarayan PR, Saini JR, Parmar DK (2015) Applications of pattern recognition algorithms in agriculture: a review. Int J Adv Netw Appl 6(5):2495–2502

Sabeena J, Venkata SRP (2019) A modified deep learning enthused adversarial network model to predict financial fluctuations in stock market. Int J Eng Adv Technol 8:2996–3000

Saravanan V, Charanya SK (2018) E-Commerce Product Classification using Lexical Based Hybrid Feature Extraction and SVM. Int J Innov Technol Explor Eng 9(1):1885–1891

Schmidhuber J (2015) Deep learning in neural networks: An overview. Neural Networks 61:85–117

Simester D, Timoshenko A, Zoumpoulis SI (2020) Targeting prospective customers: robustness of machine-learning methods to typical data challenges. Manag Sci 66(6):2495–2522

Singh R, Srivastava S (2017) Stock prediction using deep learning. Multimed Tools Appl 76(18):18569–18584

Sirignano J, Cont R (2019) Universal features of price formation in financial markets: perspectives from deep learning. Quant Financ 19(9):1449–1459

Sohangir S, Wang DD, Pomerants A, Khoshgoftaar TM (2018) Big data: deep learning for financial sentiment analysis. J Big Data 5(1):25

Song Y, Lee JW, Lee J (2019) A study on novel filtering and relationship between input-features and target-vectors in a deep learning model for stock price prediction. Appl Intell 49(3):897–911

Storm H, Baylis K, Heckelei T (2020) Machine learning in agricultural and applied economics. Eur Rev Agric Econ 47(3):849–892

Tamura K, Uenoyama K, Iitsuka S, Matsuo Y (2018) Model for evaluation of stock values by ensemble model using deep learning. Trans Jpn Soc Artif Intell 2018:33

Tashiro D, Matsushima H, Izumi K, Sakaji H (2019) Encoding of high-frequency order information and prediction of short-term stock price by deep learning. Quant Financ 19(9):1499–1506

Tibshirani R (1996) Regression shrinkage and selection via the Lasso. J R Stat Soc Ser B Stat Methodol 58(1):267–288

Timoshenko A, Hauser JR (2019) Identifying customer needs from user-generated content. Mark Sci 38(1):1–20

Trandafili E, Biba M (2013) A review of machine learning and data mining approaches for business applications in social networks. Int J E Bus Res (IJEBR) 9(1):36–53

Valencia F, Gomez-Espinosa A, Valdes-Aguirre B (2019) Price movement prediction of cryptocurrencies using sentiment analysis and machine learning. Entropy 21(6):12

Vapnik V (2013) The nature of statistical learning theory. Springer, Berlin

Vo NNY, He X, Liu S, Xu, G (2019) Deep learning for decision making and the optimization of socially responsible invest-ments and portfolio.   Decis Support Syst 124:113097. https://doi.org/10.1016/j.dss.2019.113097

Wang XY, Luo DK, Zhao X, Sun Z (2018b) Estimates of energy consumption in China using a self-adaptive multi-verse optimizer-based support vector machine with rolling cross-validation. Energy 152:539–548

Wang Y, Mo DY, Tseng MM (2018c) Mapping customer needs to design parameters in the front end of product design by applying deep learning. CIRP Ann 67(1):145–148

Wang B, Ning LJ, Kong Y (2019) Integration of unsupervised and supervised machine learning algorithms for credit risk assessment. Expert Syst Appl 128:301–315

Wang WY, Li WZ, Zhang N, Liu KC (2020) Portfolio formation with preselection using deep learning from long-term financial data. Expert Syst Appl 143:17

Wang C, Zhu H, Hu R, Li R, Jiang C (2023) LongArms: fraud prediction in online lending services using sparse knowledge graph. IEEE Trans Big Data 9(2):758–772

Wang Q, Li BB, Singh PV (2018) Copycats vs. original mobile apps: a machine learning copycat-detection method and empirical analysis. Inf Syst Res 29(2):273–291

Weng B, Lu L, Wang X, Megahed FM, Martinez W (2018) Predicting short-term stock prices using ensemble methods and online data sources. Expert Syst Appl 112:258–273

Wettschereck D, Aha DW, Mohri T (1997) A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms. Artif Intell Rev 11(1–5):273–314

Wu WB, Chen JQ, Yang ZB, Tindall ML (2021) A cross-sectional machine learning approach for hedge fund return predic-tion and selection. Manage Sci 67(7):4577–4601

Wu X, Kumar V, Ross Quinlan J, Ghosh J, Yang Q, Motoda H, McLachlan GJ, Ng A, Liu B, Yu PS, Zhou ZH, Steinbach M, Hand DJ, Steinberg D (2008) Top 10 algorithms in data mining. Knowl Inform Syst 14:1–37

Wu C, Yan M (2018) Session-aware Information Embedding for E-commerce Product Recommendation. In: Paper presented at the Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, Singapore, Singapore

Xiao F, Ke J (2021) Pricing, management and decision-making of financial markets with artificial intelligence: introduction to the issue. Financ Innov 7(1):85

Xu YZ, Zhang JL, Hua Y, Wang LY (2019) Dynamic credit risk evaluation method for e-commerce sellers based on a hybrid artificial intelligence model. Sustainability 11:5521

Xu WJ, Chen X, Dong YC, Chiclana F (2021) Impact of decision rules and non-cooperative behaviors on minimum con-sensus cost in group decision making. Group Decis Negot 30(6):1239–1260

Yan HJ, Ouyang HB (2018) Financial time series prediction based on deep learning. Wirel Pers Commun 102(2):683–700

Yao LY, Chu ZX, Li S, Li YL, Gao J, Zhang AD (2021) A survey on causal inference. ACM Trans Knowl Discov Data 15(5):1–46

Yoganarasimhan H (2020) Search personalization using machine learning. Manag Sci 66(3):1045–1070

Young D, Poletti S, Browne O (2014) Can agent-based models forecast spot prices in electricity markets? Evidence from the New Zealand electricity market. Energy Econ 45:419–434

Zahavi JN, Levin I (1997) Applying neural computing to target marketing. J Direct Mark 11(4):76–93

Zha QB, Kou G, Zhang HJ, Liang HM, Chen X, Li CC, Dong YC (2020) Opinion dynamics in finance and business: a literature review and research opportunities. Financ Innov 6(1):44

Zha QB, Dong YC, Zhang HJ, Chiclana F, Herrera-Viedma E (2021) A personalized feedback mechanism based on bounded confidence learning to support consensus reaching in group decision making. IEEE Trans Syst Man Cybern Syst 51(6):3900–3910

Zhang QG, Benveniste A (1992) Wavelet networks. IEEE Trans Neural Netw 3(6):889–898

Zhang C, Li R, Shi H, Li FR (2020a) Deep learning for day-ahead electricity price forecasting. IET Smart Grid 3(4):462–469

Zhang YJ, Li BB, Krishnan R (2020b) Learning Individual behavior using sensor data: the case of global positioning system traces and taxi drivers. Inf Syst Res 31(4):1301–1321

Zhang B, Tan RH, Lin CJ (2021a) Forecasting of e-commerce transaction volume using a hybrid of extreme learning machine and improved moth-flame optimization algorithm. Appl Intell 51(2):952–965

Zhang HJ, Li CC, Liu YT, Dong YC (2021b) Modelling personalized individual semantics and consensus in comparative linguistic expression preference relations with self-confidence: An optimization-based approach. IEEE Trans Fuzzy Syst 29:627–640

Zhao L (2021) The function and impact of cryptocurrency and data technology in the context of financial technology: introduction to the issue. Financ Innov 7(1):84

Zhu XD, Ninh A, Zhao H, Liu ZM (2021) Demand forecasting with supply-chain information and machine learning: evidence in the pharmaceutical industry. Prod Oper Manag 30(9):3231–3252

Publisher's Note