REVIEW   **Open Access**

Deep learning in finance and banking: A literature review and classification

Jian Huang[1], Junyi Chai[2]* and Stella Cho[2]

* Correspondence: donchaiam@
gmail.com; donjychai@uic.edu.hk
[2]Division of Business and
Management, BNU-HKBU United
International College, Zhuhai, China
Full list of author information is
available at the end of the article

Abstract

Deep learning has been widely applied in computer vision, natural language processing, and audio-visual recognition. The overwhelming success of deep learning as a data processing technique has sparked the interest of the research community. Given the proliferation of Fintech in recent years, the use of deep learning in finance and banking services has become prevalent. However, a detailed survey of the applications of deep learning in finance and banking is lacking in the existing literature. This study surveys and analyzes the literature on the application of deep learning models in the key finance and banking domains to provide a systematic evaluation of the model preprocessing, input data, and model evaluation. Finally, we discuss three aspects that could affect the outcomes of financial deep learning models. This study provides academics and practitioners with insight and direction on the state-of-the-art of the application of deep learning models in finance and banking.

Keywords: Literature review, Deep learning, Finance, Banking, Fintech

Introduction

Deep learning (DL) is an advanced technique of machine learning (ML) based on artificial neural network (NN) algorithms. As a promising branch of artificial intelligence, DL has attracted great attention in recent years. Compared with conventional ML techniques such as support vector machine (SVM) and k-nearest neighbors (kNN), DL possesses advantages of the unsupervised feature learning, a strong capability of generalization, and a robust training power for big data. Currently, DL has been applied comprehensively in classification and prediction tasks, computer visions, image processing, and audio-visual recognition (Chai and Li 2019). Although DL was developed in the field of computer science, its applications have penetrated diversified fields such as medicine, neuroscience, physics and astronomy, finance and banking (F&B), and operations management (Chai et al. 2013; Chai and Ngai 2020). The existing literature lacks a good overview of DL applications in F&B fields. This study attempts to bridge this gap.

While DL is the focus of computer vision (e.g., Elad and Aharon 2006; Guo et al. 2016) and natural language processing (e.g., Collobert et al. 2011) in the mainstream,

DL applications in F&B are developing rapidly. Shravan and Vadlamani (2016) investigated the tools of text mining for F&B domains. They examined the representative ML algorithms, including SVM, kNN, genetic algorithm (GA), and AdaBoost. Butaru et al. (2016) compared performances of DL algorithms, including random forests, decision trees, and regularized logistic regression. They found that random forests gained the highest classification accuracy in the delinquency status.

Cavalcante et al. (2016) summarized the literature published from 2009 to 2015. They analyzed DL models, including multi-layer perceptron (MLP) (a fast library for approximate nearest neighbors), Chebyshev functional link artificial NN, and adaptive weighting NN. Although the study constructed a prediction framework in financial trading, some notable DL techniques such as long short-term memory (LSTM) and reinforcement learning (RL) models are neglect. Thus, the framework cannot ascertain the optimal model in a specific condition.

The reviews of the existing literature are either incomplete or outdated. However, our study provides a comprehensive and state-of-the-art review that could capture the relationships between typical DL models and various F&B domains. We identified critical conditions to limit our collection of articles. We employed academic databases in Science Direct, Springer-Link Journal, IEEE Xplore, Emerald, JSTOR, ProQuest Database, EBSCOhost Research Databases, Academic Search Premier, World Scientific Net, and Google Scholar to search for articles. We used two groups of keywords for our search. One group is related to the DL, including "deep learning," "neural network," "convolutional neural networks" (CNN), "recurrent neural network" (RNN), "LSTM," and "RL." The other group is related to finance, including "finance," "market risk," "stock risk," "credit risk," "stock market," and "banking." It is important to conduct cross searches between computer-science-related and finance-related literature. Our survey exclusively focuses on the financial application of DL models rather than other DL models like SVM, kNN, or random forest. The time range of our review was set between 2014 and 2018. In this stage, we collected more than 150 articles after cross-searching. We carefully reviewd each article and considered whether it is worthy of entering our pool of articles for review. We removed the articles if they are not from reputable journals or top professional conferences. Moreover, articles were discarded if the details of financial DL models presented were not clarified. Thus, 40 articles were selected for this review eventually.

This study contributes to the literature in the following ways. First, we systematically review the state-of-the-art applications of DL in F&B fields. Second, we summarize multiple DL models regarding specified F&B domains and identify the optimal DL model of various application scenarios. Our analyses rely on the data processing methods of DL models, including preprocessing, input data, and evaluation rules. Third, our review attempts to bridge the technological and application levels of DL and F&B, respectively. We recognize the features of various DL models and highlight their feasibility toward different F&B domains. The penetration of DL into F&B is an emerging trend. Researchers and financial analysts should know the feasibilities of particular DL models toward a specified financial domain. They usually face difficulties due to the lack of connections between core financial domains and numerous DL models. This study will fill this literature gap and guide financial analysts.

The rest of this paper is organized as follows. Section 2 provides a background of DL techniques. Section 3 introduces our research framework and methodology. Section 4

analyzes the established DL models. Section 5 analyzes key methods of data processing, including data preprocessing and data inputs. Section 6 captures appeared criteria for evaluating the performance of DL models. Section 7 provides a general comparison of DL models against identified F&B domains. Section 8 discusses the influencing factors in the performance of financial DL models. Section 9 concludes and outlines the scope for promising future studies.

Background of deep learning

Regarding DL, the term "deep" presents the multiple layers that exist in the network. The history of DL can be traced back to stochastic gradient descent in 1952, which is employed for an optimization problem. The bottleneck of DL at that time was the limit of computer hardware, as it was very time-consuming for computers to process the data. Today, DL is booming with the developments of graphics processing units (GPUs), dataset storage and processing, distributed systems, and software such as Tensor Flow. This section briefly reviews the basic concept of DL, including NN and deep neural network (DNN). All of these models have greatly contributed to the applications in F&B.

The basic structure of NN can be illustrated as $Y = F(X^T w + c)$ regarding the independent (input) variables X, the weight terms w, and the constant terms c. Y is the dependent variable and X is formed as an $n \times m$ matrix for the number of training sample n and the number of input variables m. To apply this structure in finance, Y can be considered as the price of next term, the credit risk level of clients, or the return rate of a portfolio. F is an activation function that is unique and different from regression models. F is usually formulated as sigmoid functions and tanh functions. Other functions can also be used, including ReLU functions, identity functions, binary step functions, ArcTan functions, ArcSinh functions, ISRU functions, ISRLU functions, and SQNL functions. If we combine several perceptrons in each layer and add a hidden layer from Z_1 to Z_4 in the middle, we term a single layer as a neural network, where the input layers are the X_s, and the output layers are the Y_s. In finance, Y can be considered as the stock price. Moreover, multiple Y_s are also applicable; for instance, fund managers often care about future prices and fluctuations. Figure 1 illustrates the basic structure.

Based on the basic structure of NN shown in Fig. 1, traditional networks include DNN, backpropagation (BP), MLP, and feedforward neural network (FNN). Using these models can ignore the order of data and the significance of time. As shown in Fig. 2, RNN has a new NN structure that can address the issues of long-term dependence and the order between input variables. As financial data in time series are very common, uncovering hidden correlations is critical in the real world. RNN can be better at solving this problem, as compared to other moving average (MA) methods that have been frequently adopted before. A detailed structure of RNN for a sequence over time is shown in Part B of the Appendix (see Fig. 7 in Appendix).

Although RNN can resolve the issue of time-series order, the issue of long-term dependencies remains. It is difficult to find the optimal weight for long-term data. LSTM, as a type of RNN, added a gated cell to overcome long-term dependencies by combining different activation functions (e.g., sigmoid or tanh). Given that LSTM is frequently used for forecasting in the finance literature, we extract LSTM from RNN models and name other structures of standard RNN as RNN(O).

Fig. 1 The structure of NN
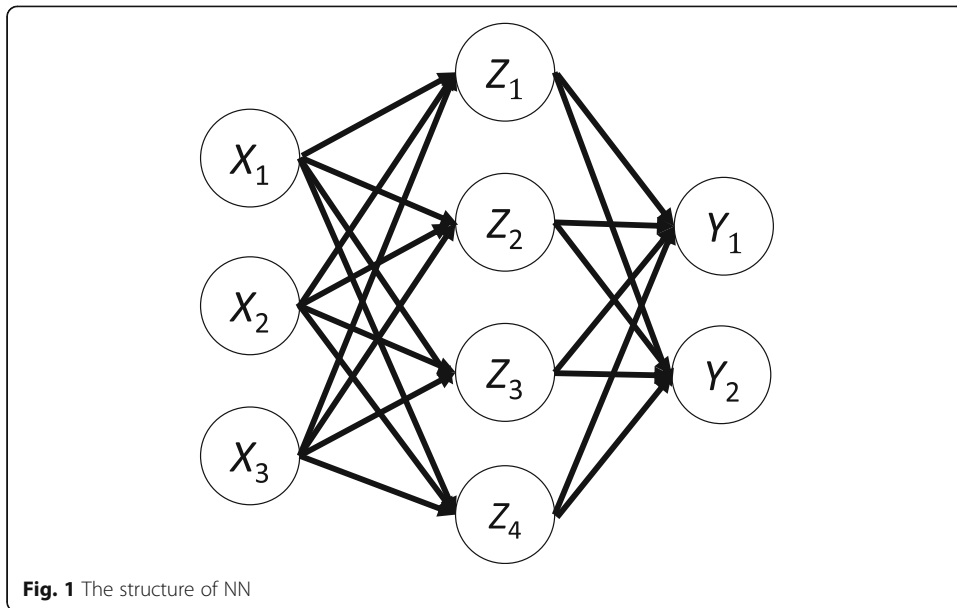
As we focus on the application rather than theoretical DL aspect, this study will not consider other popular DL algorithms, including CNN and RL, as well as Latent variable models such as variational autoencoders and generative adversarial network. Table 6 in Appendix shows a legend note to explain the abbreviations used in this paper. We summarize the relationship between commonly used DL models in Fig. 3.

## Research framework and methodology

Our research framework is illustrated in Fig. 4. We combine qualitative and quantitative analyses of the articles in this study. Based on our review, we recognize and identify seven core F&B domains, as shown in Fig. 5. To connect the DL side and the F&B side,

Fig. 2 The abstract structure of RNN

**Fig. 3** Relationships of reviewed DL models for F&B domains
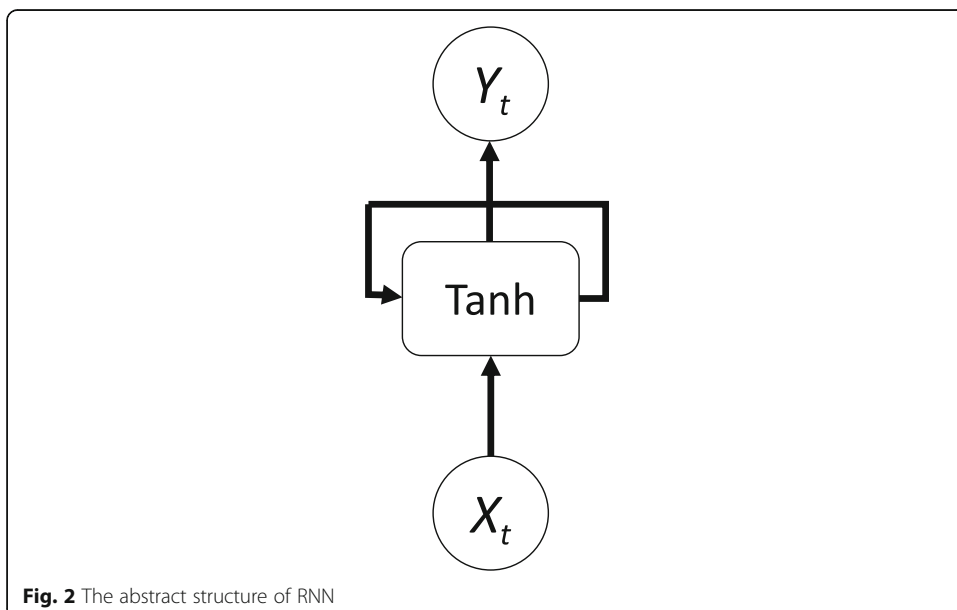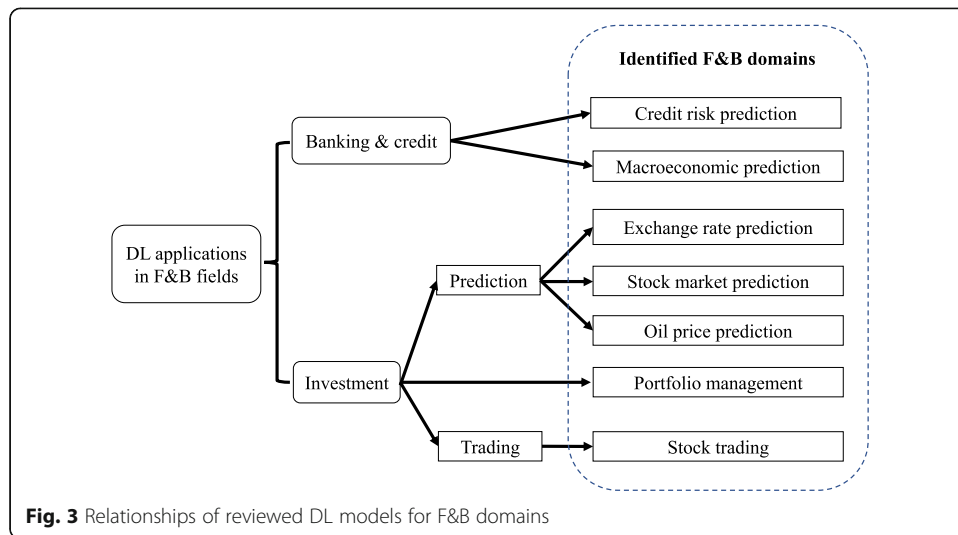
we present our review on the application of the DL model in seven F&B domains in Section 4. It is crucial to analyze the feasibility of a DL model toward particular domains. To do so, we provide summarizations in three key aspects, including data preprocessing, data inputs, and evaluation rules, according to our collection of articles. Finally, we determine optimal DL models regarding the identified domains. We further discuss two common issues in using DL models for F&B: overfitting and sustainability.

Figure 5 shows that the application domains can be divided into two major areas: (1) banking and credit risk and (2) financial market investment. The former contains two domains: credit risk prediction and macroeconomic prediction. The latter contains

Fig. 4 The research framework of this study

Fig. 5 The identified domains of F&B for DL applications
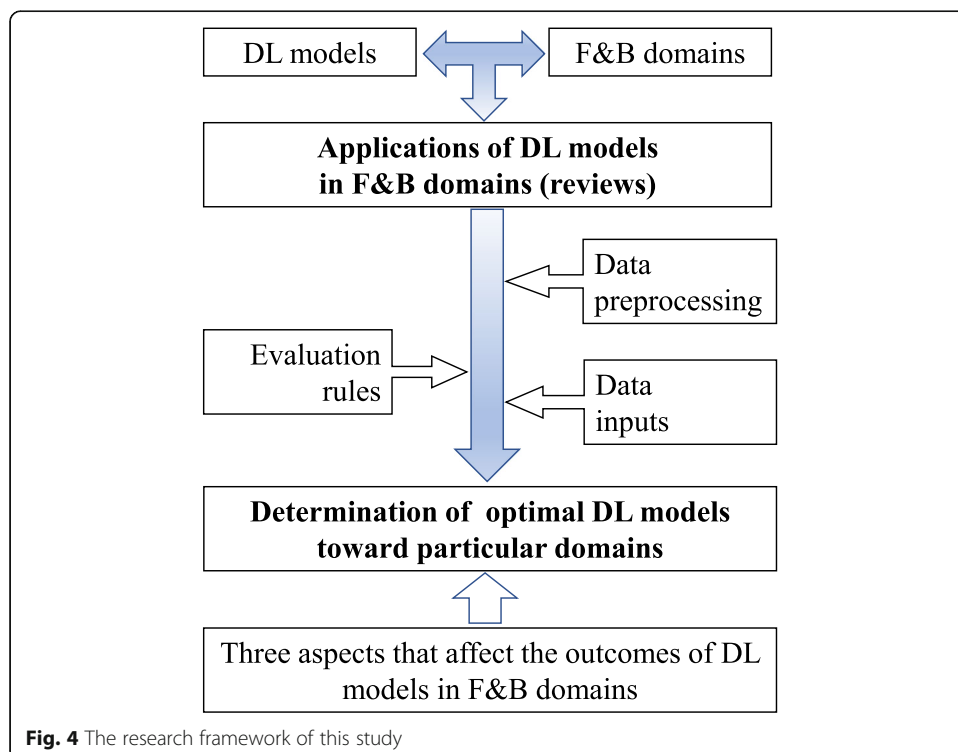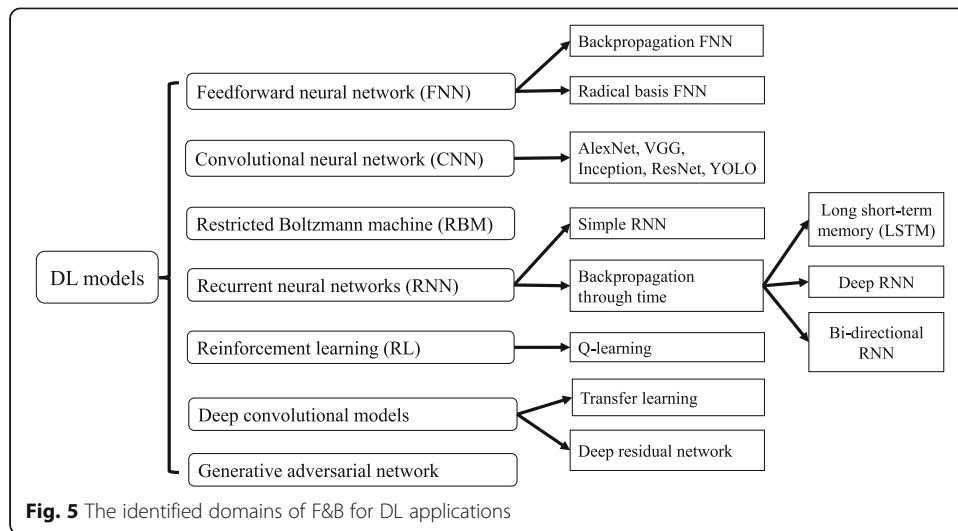
financial prediction, trading, and portfolio management. Prediction tasks are crucial, as emphasized by Cavalcante et al. (2016). We study this domain from three aspects of prediction, including exchange rate, stock market, and oil price. We illustrate this structure of application domains in F&B.

Figure 6 shows a statistic in the listed F&B domains. We illustrate the domains of financial applications on the X-axis and count the number of articles on the Y-axis. Note that a reviewed article could cover more than one domain in this figure; thus, the sum of the counts (45) is larger than the size of our review pool (40 articles). As shown in Fig. 6, stock marketing prediction and trading dominate the listed domains, followed by exchange rate prediction. Moreover, we found two articles on banking credit risk and two articles on portfolio management. Price prediction and macroeconomic prediction are two potential topics that deserve more studies.

Fig. 6 A count of articles over seven identified F&B domains
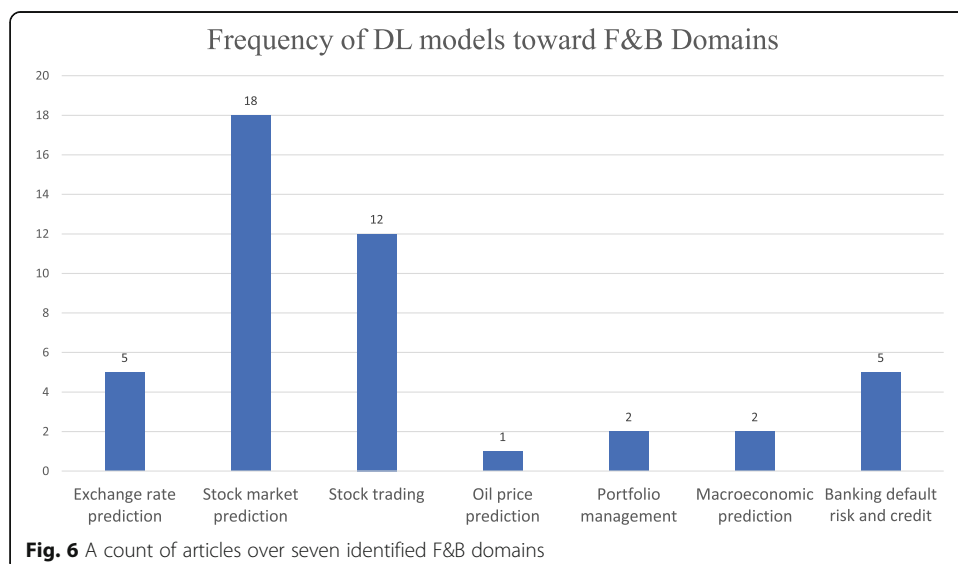
Application of DL models in F&B domains

Based on our review, six types of DL models are reported. They are FNN, CNN, RNN, RL, deep belief networks (DBN), and restricted Boltzmann machine (RBM). Regarding FNN, several papers use the alternative terms of backpropagation artificial neural network (ANN), FNN, MLP, and DNN. They have an identical structure. Regarding RNN, one of its well-known models in the time-series analysis is called LSTM. Nearly half of the reviewed articles apply FNN as the primary DL technique. Nine articles apply LSTM, followed by eight articles for RL, and six articles for RNN. Minor ones that are applied in F&B include CNN, DBM, and RBM. We count the number of articles that use various DL models in seven F&B domains, as shown in Table 1. FNN is the principal model used in exchange rate, price, and macroeconomic predictions, as well as banking default risk and credit. LSTM and FNN are two kinds of popular models for stock market prediction. Differently, RL and FNN are frequently used regarding stock trading. FNN, RL, and simple RNN can be conducted in portfolio management. FNN is the primary model in macroeconomic and banking risk prediction. CNN, LSTM, and RL are emerging research approaches in banking risk prediction. The detailed statistics that contain specific articles can be found in Table 5 in Appendix.

Exchange rate prediction

Shen et al. (2015) construct an improved DBN model by including RBM and find that their model outperforms the random walk algorithm, auto-regressive-moving-average (ARMA), and FNN with fewer errors. Zheng et al. (2017) examine the performance of DBN and find that the DBN model estimates the exchange rate better than FNN model does. They find that a small number of layer nodes engender a more significant effect on DBN.

Several scholars believe that a hybrid model should have better performance. Ravi et al. (2017) contribute a hybrid model by using MLP (FNN), chaos theory, and multi-objective evolutionary algorithms. Their Chaos+MLP + NSGA-II model[1] has a mean squared error (MSE) with 2.16E-08 that is very low. Several articles point out that only a complicated neural network like CNN can gain higher accuracy. For example, Galeshchuk and Mukherjee (2017) conduct experiments and claim that a single hidden layer NN or SVM performs worse than a simple model like moving average (MA). However, they find that CNN could achieve higher classification accuracy in predicting the direction of the change of exchange rate because of successive layers of DNN.

Stock market prediction

In stock market prediction, some studies suggest that market news may influence the stock price and DL model, such as using a magic filter to extract useful information for price prediction. Matsubara et al. (2018) extract information from the news and propose a deep neural generative model to predict the movement of the stock price. This model combines DNN and a generative model. It suggests that this hybrid approach outperforms SVM and MLP.

Minh et al. (2017) develop a novel framework with two streams combining the gated recurrent unit network and the Stock2vec. It employs a word embedding and sentiment

---

[1]In the model, NSGA stands for non-dominated sorting genetic algorithm.

Table 1 Statistics of DL models in seven domains of financial applications

		ANN/DNN/FNN/MLP	CNN	LSTM	RNN(O)	RL	DBN	RBM	Total
1	Exchange rate prediction	3	1	0	0	0	2	0	6
2	Stock market prediction	6	3	7	4	0	0	2	22
3	Stock trading	6	1	3	3	6	0	0	19
4	Banking default risk and credit	3	1	1	0	1	0	0	6
5	Portfolio management	1	0	0	1	1	0	0	3
6	Macroeconomic prediction	2	0	0	0	0	0	0	2
7	Oil price prediction	1	0	0	0	0	0	0	1
Total		22	6	11	8	8	2	2	59

training system on financial news and the Harvard IV-4 dataset. They use the historical price and news-based signals from the model to predict the S&P500 and VN-index price directions. Their model shows that the two-stream gated recurrent unit is better than the gated recurrent unit or the LSTM. Jiang et al. (2018) establish a recurrent NN that extracts the interaction between the inner-domain and cross-domain of financial information. They prove that their model outperforms the simple RNN and MLP in the currency and stock market. Krausa and Feuerriegel (2017) propose that they can transform financial disclosure into a decision through the DL model. After training and testing, they point out that LSTM works better than the RNN and conventional ML methods such as ridge regression, Lasso, elastic net, random forest, SVR, AdaBoost, and gradient boosting. They further pre-train words embeddings with transfer learning (Krausa and Feuerriegel 2017). They conclude that better performance comes from LSTM with word embeddings. In the sentiment analysis, Sohangir et al. (2018) compares LSTM, doc2vec, and CNN to evaluate the stock opinions on the StockTwits. They conclude that CNN is the optimal model to predict the sentiment of authors. This result may be further applied to predict the stock market trend.

Data preprocessing is conducted to input data into the NN. Researchers may apply numeric unsupervised methods of feature extraction, including principal component analysis, autoencoder, RBM, and kNN. These methods can reduce the computational complexity and prevent overfitting. After the input of high-frequency transaction data, Chen et al. (2018b) establish a DL model with an auto-encoder and an RBM. They compare their model with backpropagation FNN, extreme learning machine, and radial basis FNN. They claim that their model can better predict the Chinese stock market. Chong et al. (2017) apply the principal component analysis (PCA) and RBM with high-frequency data of the South Korean market. They find that their model can explain the residual of the autoregressive model. The DL model can thus extract additional information and improve prediction performance. More so, Singh and Srivastava (2017) describe a model involving 2-directional and 2-dimensional ($2D^2$) PCA and DNN. Their model outperforms $2D^2$ with radial basis FNN and RNN.

For time-series data, sometimes it is difficult to judge the weight of long-term and short-term data. The LSTM model is just for resolving this problem in financial prediction. The literature has attempted to prove that LSTM models are applicable and

outperform conventional FNN models. Yan and Ouyang (2017) apply LSTM to challenge the MLP, SVM, and kNN in predicting a static and dynamic trend. After a wavelet decomposition and a reconstruction of the financial time series, their model can be used to predict a long-term dynamic trend. Baek and Kim (2018) apply LSTM not only in predicting the price of S&P500 and KOSPI200 but also in preventing overfitting. Kim and Won (2018) apply LSTM in the prediction of stock price volatility. They propose a hybrid model that combines LSTM with three generalized autoregressive conditional heteroscedasticity (GARCH)-type models. Hernandez and Abad (2018) argue that RBM is inappropriate for dynamic data modeling in the timeseries analysis because it cannot retain memory. They apply a modified RBM model called p-RBM that can retain the memory of p past states. This model is used in predicting market directions of the NASDAQ-100 index. Compared with vector autoregression (VAR) and LSTM, notwithstanding, they find that LSTM is better because it can uncover the hidden structure within the non-linear data while VAR and p-RBM cannot capture the non-linearity in data.

CNN was established to predict the price with a complicated structure. Making the best use of historical price, Dingli and Fournier (2017) develop a new CNN model. This model can predict next month's price. Their results cannot surpass other comparable models, such as logistic regression (LR) and SVM. Tadaaki (2018) applies the financial ratio and converts them into a "grayscale image" in the CNN model. The results reveal that CNN is more efficient than decision trees (DT), SVM, linear discriminant analysis, MLP, and AdaBoost. To predict the stock direction, Gunduz et al. (2017) establish a CNN model with a so-called specially ordered feature set whose classifier outperforms either CNN or LR.

Stock trading

Many studies adopt the conventional FNN model and try to set up a profitable trading system. Sezer et al. (2017) combine GA with MLP. Chen et al. (2017) adopt a double-layer NN and discover that its accuracy is better than ARMA-GARCH and single-layer NN. Hsu et al. (2018) equip the Black-Scholes model and a three-layer fully-connected feedforward network to estimate the bid-ask spread of option price. They argue that this novel model is better than the conventional Black-Scholes model with lower RMSE. Krauss et al. (2017) apply DNN, gradient-boosted-trees, and random forests in statistical arbitrage. They argue that their returns outperform the market index S&P500.

Several studies report that RNN and its derivate models are potential. Deng et al. (2017) extend the fuzzy learning into the RNN model. After comparing their model to different DL models like CNN, RNN, and LSTM, they claim that their model is the optimal one. Fischer and Krauss (2017) and Bao et al. (2017) argue that LSTM can create an optimal trading system. Fischer and Krauss (2017) claim that their model has a daily return of 0.46 and a sharp ratio of 5.8 prior to the transaction cost. Given the transaction cost, however, LSTM's profitability fluctuated around zero after 2010. Bao et al. (2017) advance Fischer and Krauss's (2017) work and propose a novel DL model (i.e., WSAEs-LSTM model). It uses wavelet transforms to eliminate noise, stacked autoencoders (SAEs) to predict stock price, and LSTM to predict the close price. The result

shows that their model outperforms other models such as WLSTM,[2] LSTM, and RNN in predictive accuracy and profitability.

RL is popular recently despite its complexity. We find that five studies apply this model. Chen et al. (2018a) propose an agent-based RL system to mimic 80% professional trading strategies. Feuerriegel and Prendinger (2016) convert the news sentiment into the signal in the trading system, although their daily returns and abnormal returns are nearly zero. Chakraborty (2019) cast the general financial market fluctuation into a stochastic control problem and explore the power of two RL models, including Q-learning[3] and state-action-reward-state-action (SARSA) algorithm. Both models can enhance profitability (e.g., 9.76% for Q-learning and 8.52% for SARSA). They outperform the buy-and-hold strategy.[4] Zhang and Maringer (2015) conduct a hybrid model called GA, with recurrent RL. GA is used to select an optimal combination of technical indicators, fundamental indicators, and volatility indicators. The out-of-sample trading performance is improved due to a significantly positive Sharpe ratio. Martinez-Miranda et al. (2016) create a new topic of trading. It uses a market manipulation scanner model rather than a trading system. They use RL to model spoofing-and-pinging trading. This study reveals that their model just works on the bull market. Jeong and Kim (2018) propose a model called deep Q-network that is constructed by RL, DNN, and transfer learning. They use transfer learning to solve the overfitting issue incurred as a result of insufficient data. They argue that the profit yields in this system increase by four times the amount in S&P500, five times in KOSPI, six times in EuroStoxx50, and 12 times in HIS.

Banking default risk and credit

Most articles in this domain focus on FNN applications. Rönnqvist and Sarlin (2017) propose a model for detecting relevant discussions in texting and extracting natural language descriptions of events. They convert the news into a signal of the bank-distress report. In their back-test, their model reflects the distressing financial event of the 2007–2008 period.

Zhu et al. (2018) propose a hybrid CNN model with a feature selection algorithm. Their model outperforms LR and random forest in consumer credit scoring. Wang et al. (2019) consider that online operation data can be used to predict consumer credit scores. They thus convert each kind of event into a word and apply the Event2vec model to transform the word into a vector in the LSTM network. The probability of default yields higher accuracy than other models. Jurgovsky et al. (2018) employs the LSTM to detect credit card fraud and find that LSTM can enhance detection accuracy.

Han et al. (2018) report a method that adopts RL to assess the credit risk. They claim that high-dimensional partial differential equations (PDEs) can be reformulated by using backward stochastic differential equations. NN approximates the gradient of the unknown solution. This model can be applied to F&B risk evaluation after considering all elements such as participating agents, assets, and resources, simultaneously.

---

[2]A combination of Wavelet transforms (WT) and long-short term memory (LSTM) is called WLSTM in Bao et al. (2017).

[3]Q-learning is a model-free reinforcement learning algorithm.

[4]Buy-and-hold is a passive investment strategy in which an investor buys stocks (or ETFs) and holds them for a long period regardless of fluctuations in the market.

Portfolio management

Song et al. (2017) establish a model after combining ListNet and RankNet to make a portfolio. They take a long position for the top 25% stocks and hold the short position for the bottom 25% stocks weekly. The ListNetlong-short model is the optimal one, which can achieve a return of 9.56%. Almahdi and Yang (2017) establish a better portfolio with a combination of RNN and RL. The result shows that the proposed trading system respond to transaction cost effects efficiently and outperform hedge fund benchmarks consistently.

Macroeconomic prediction

Sevim et al. (2014) develops a model with a back-propagation learning algorithm to predict the financial crises up to a year before it happened. This model contains three-layer perceptrons (i.e., MLP) and can achieve an accuracy rate of approximately 95%, which is superior to DT and LR. Chatzis et al. (2018) examine multiple models such as classification tree, SVM, random forests, DNN, and extreme gradient boosting to predict the market crisis. The results show that crises encourage persistence. Furthermore, using DNN increases the classification accuracy that makes global warning systems more efficient.

Price prediction

For price prediction, Sehgal and Pandey (2015) review ANN, SVM, wavelet, GA, and hybrid systems. They separate the time-series models into stochastic models, AI-based models, and regression models to predict oil prices. They reveal that researchers prevalently use MLP for price prediction.

Data preprocessing and data input

Data preprocessing

Data preprocessing is conducted to denoise before data training of DL. This section summarizes the methods of data preprocessing. Multiple preprocessing techniques discussed in Part 4 include the principal component analysis (Chong et al. 2017), SVM (Gunduz et al. 2017), autoencoder, and RBM (Chen et al. 2018b). There are several additional techniques of feature selection as follows.

(1) Relief: The relief algorithm (Zhu et al. 2018) is a simple approach to weigh the importance of the feature. Based on NN algorithms, relief repeats the process for n times and divides each final weight vector by n. Thus, the weight vectors are the relevance vectors, and features are selected if their relevance is larger than the threshold τ.

(2) Wavelet transforms: Wavelet transforms are used to fix the noise feature of the financial time series before feeding into a DL network. It is a widely used technique for filtering and mining single-dimensional signals (Bao et al. 2017).

(3) Chi-square: Chi-square selection is commonly used in ML to measure the dependence between a feature and a class label. The representative usage is by Gunduz et al. (2017).

(4) Random forest: Random forest algorithm is a two-stage process that contains random feature selection and bagging. The representative usage is by Fischer and Krauss (2017).

Data inputs

Data inputs are an important criterion for judging whether a DL model is feasible for particular F&B domains. This section summarizes the method of data inputs that have been adopted in the literature. Based on our review, five types of input data in the F&B domain can be presented. Table 2 provides a detailed summary of the input variable in F&B domains.

(1) History price: The daily exchange rate can be considered as history price. The price can be the high, low, open, and close price of the stock. Related articles include Bao et al. (2017), Chen et al. (2017), Singh and Srivastava (2017), and Yan and Ouyang (2017).

(2) Technical index: Technical indexes include MA, exponential MA, MA convergence divergence, and relative strength index. Related articles include Bao et al. (2017), Chen et al. (2017), Gunduz et al. (2017), Sezer et al. (2017), Singh and Srivastava (2017), and Yan and Ouyang (2017).

(3) Financial news: Financial news covers financial message, sentiment shock score, and sentiment trend score. Related articles include Feuerriegel and Prendinger (2016), Krausa and Feuerriegel (2017), Minh et al. (2017), and Song et al. (2017).

(4) Financial report data: Financial report data can account for items in the financial balance sheet or the financial report data (e.g., return on equity, return on assets, price to earnings ratio, and debt to equity ratio). Zhang and Maringer (2015) is a representative study on the subject.

(5) Macroeconomic data: This kind of data includes macroeconomic variables. It may affect elements of the financial market, such as exchange rate, interest rate, overnight interest rate, and gross foreign exchange reserves of the central bank. Representative articles include Bao et al. (2017), Kim and Won (2018), and Sevim et al. (2014).

(6) Stochastic data: Chakraborty (2019) provides a representative implementation.

Evaluation rules

It is critical to judge whether an adopted DL model works well in a particular financial domain. We, thus, need to consider evaluation systems of criteria for gauging the performance of a DL model. This section summarizes the evaluation rules of F&B-oriented DL models. Based on our review, three evaluation rules dominate: the error term, the accuracy index, and the financial index. Table 3 provides a detailed summary. The evaluation rules can be boiled down to the following categories.

(1) Error term: Suppose Y_{t+i} and F_{t+i} are the real data and the prediction data, respectively, where m is the total number. The following is a summary of the functional formula commonly employed for evaluating DL models.

Table 2 A summary of the input variables of our reviewed articles in F&B domains

	Publication	Input variable
Exchange rate prediction	Shen et al. (2015)	Daily Exchange rate
	Zheng et al. (2017)	Daily Exchange rate
	Ravi et al. (2017)	Daily Exchange rate
	Galeshchuk and Mukherjee (2017)	Daily Exchange rate
Stock market prediction	Matsubara et al. (2018)	Technical index, specific events (corporate buyouts, product releases)
	Chen et al. (2018b)	1-min transaction data
	Chong et al. (2017)	Daily and five-minute return
	Yan and Ouyang (2017)	Volume, technical index, Price (Open, Close, High, Low, Volume)
	Kim and Won (2018)	GARCH, EGARCH, KOSPI 200 Index, gold price, CB interest rate, KTB interest rate, KOSPI200 index log difference
	Singh and Srivastava (2017)	Technical index, Price (Open, Close, Low, Close)
	Dingli and Fournier (2017)	Daily data, weekly data, monthly data
	Gunduz et al. (2017)	Technical indicators, price and temporal information, Chi-square feature selection
	Hernandez and Abad (2018)	Daily data
	Sohangir et al. (2018)	Message ID, a user ID, the author's number of followers, a timestamp, the current price of the stock, and other record-keeping attributes
	Krausa and Feuerriegel (2017)	Financial news, stock market data
	Minh et al. (2017)	Financial news, daily stock price
Stock trading	Sezer et al. (2017)	Close/high price,technical index (RSI, SMA)
	Chen et al. (2017)	Technical indicators
	Deng et al. (2017)	Minute-level close prices, daily price
	Fischer and Krauss (2017)	Daily price
	Bao et al. (2017)	Daily trading data (open/close/high/low, trading volume), technical indicator, macroeconomic variable (exchange rate, interest rate)
	Martinez-Miranda et al. (2016)	Daily data
	Feuerriegel and Prendinger (2016)	Financial news
	Chakraborty (2019)	Sum of discounted reward, stochastic data
	Zhang and Maringer (2015)	Daily data of 180 SandP stocks, positive volume index and negative volume index, exponential MA, relative strength index, cash flow (P–CF), price to earnings (P–E), debt to equity market (D–M), Conditional volatility (CVOL) (GARCH model to retrieve the conditional volatility)
Price prediction	Sehgal and Pandey (2015)	Daily price
Portfolio management	Song et al. (2017)	Financial news, sentiment shock score, sentiment trend score, price, (positive, negative and objective), relevance
	Almahdi and Yang (2017)	Five exchange-traded funds: IWD, IWC, SPY, DEM, CLY
Macroeconomic prediction	Sevim et al. (2014)	Dependent Variable: Financial Pressure Index (FPI) (includes percentage change in the

Table 2 A summary of the input variables of our reviewed articles in F&B domains *(Continued)*

	Publication	Input variable
		dollar exchange rate, gross foreign exchange reserves of Central bank, overnight interest rate); Indepedent Variable: 32 macroeconomic indicators
Banking default risk and credit	Rönnqvist and Sarlin (2017)	Event non-coinciding, ambiguous data
	Han et al. (2018)	Defaultable securities, higher interest rates for borrowing than for lending, transaction costs, uncertainties in the model parameters

- Mean Absolute Error (MAE): $\sum_{i=1}^{m} \frac{|Y_{t+i} - F_{t+i}|}{m}$;

- Mean Absolute Percent Error (MAPE): $\frac{100}{m} \sum_{i=1}^{m} \frac{|Y_{t+i} - F_{t+i}|}{Y_{t+i}}$;

- Mean Squared Error (MSE): $\sum_{i=1}^{m} \frac{(Y_{t+i} - F_{t+i})^2}{m}$;

- Root Mean Squared Error (RMSE): $\sqrt{\sum_{i=1}^{m} \frac{(Y_{t+i} - F_{t+i})^2}{m}}$;

- Normalized Mean Square Error (NMSE): $\frac{1}{m} \frac{\sum (Y_{t+i} - F_{t+i})^2}{var(Y_{t+i})}$.

(2) Accuracy index: According to Matsubara et al. (2018), we use TP, TN, FP, and FN to represent the number of true positives, true negatives, false positives, and false negatives, respectively, in a confusion matrix for classification evaluation. Based on our review, we summarize the accuracy indexes as follows.

- Directional Predictive Accuracy (DPA): $\frac{1}{N} \sum_{t=1}^{N} D_t$, if $(Y_{t+1} - Y_t) \times (F_{t+1} - Y_t) \geq 0$, $D_t = 1$, otherwise, $D_t = 0$;
- Actual Correlation Coefficient (ACC): $\frac{TP+TN}{TP+FP+FN+TN}$;
- Matthews Correlation Coefficient (MCC): $\frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$.

(3) Financial index: Financial indexes involve total return, Sharp ratio, abnormal return, annualized return, annualized number of transaction, percentage of success, average profit percent per transaction, average transaction length, maximum profit percentage in the transaction, maximum loss percentage in the transaction, maximum capital, and minimum capital.

For the prediction by regressing the numeric dependent variables (e.g., exchange rate prediction or stock market prediction), evaluation rules are mostly error terms. For the prediction by classification in the category data (e.g., direction prediction on oil price), the accuracy indexes are widely conducted. For stock trading and portfolio management, financial indexes are the final evaluation rules.

Table 3 A summary of evaluation rules for DL models in F&B domains

	Revised article	Evaluation
Exchange rate prediction	Shen et al. (2015)	RMSE, MAE, MAPE, CORR, direct accuracy
	Zheng et al. (2017)	Fitting error, MAPE
	Ravi et al. (2017)	Diebold-Mariano test, MSE, and directional change statistic
	Galeshchuk and Mukherjee (2017)	Classification accuracy
Stock market prediction	Matsubara et al. (2018)	ACC, MCC
	Chen et al. (2018b)	RMSE, MAPE, DPA
	Chong et al. (2017)	NMSE, RMSE, MAE, Mutual information
	Yan and Ouyang (2017)	MAPE, MAE
	Kim and Won (2018)	Realized volatility, loss function (MAE, MSE, HMAE, HMSE), DM, WS test
	Singh and Srivastava (2017)	Hit rate, the correlation coefficient between the actual value and prediction value, the non-linear regression multiple correlation coefficient, the correlation coefficient between the actual value and prediction return, the percentage of correct direction, the symmetric mean absolute percentage error, MAPE, RMSE, the total return
	Dingli and Fournier (2017)	Learning rate, accuracy
	Gunduz et al. (2017)	Relative MA, F-measure
	Hernandez and Abad (2018)	Means of misclassification error
	Sohangir et al. (2018)	Accuracy, F-measure, AUC, precision
	Krausa and Feuerriegel (2017)	Classification (e.g., accuracy, balanced accuracy, AUC), regression (e.g., RMSE, MSE, MAE)
	Minh et al. (2017)	NMSE, RMSE, MAE, mutual information
Stock trading	Sezer et al. (2017)	Annualized return, annualized number of transactions, percent of success, average profit percent per transaction, average transaction length, maximum profit percentage in transaction, maximum loss percentage in transaction, maximum capital, minimum capital, idle ratio
	Chen et al. (2017)	Absolute return/ Sharpe ratio/prediction accuracy
	Deng et al. (2017)	Robustness verification
	Fischer and Krauss (2017)	Accuracy, Sharp ratio, standard deviation, return
	Bao et al. (2017)	MAPE, R, Theil U
	Martinez-Miranda et al. (2016)	Average profit, STD
	Feuerriegel and Prendinger (2016)	Average daily return, Sharp ratio, abnormal return, volatility
	Chakraborty (2019)	Average annual return
	Zhang and Maringer (2015)	Exponential MA Sharpe ratio
Price prediction	Sehgal and Pandey (2015)	Prediction accuracy
Portfolio management	Song et al. (2017)	Return, volatility, Sharp ratio, maximum drawdown
	Almahdi and Yang (2017)	Coherent downside risk measure, the expected maximum drawdown, E (MDD), Calmar ratio
Macroeconomic prediction	Sevim et al. (2014)	Accuracy, sensitivity, specificity
Banking default risk and credit	Rönnqvist and Sarlin (2017)	Vector level, error, recall (precision)
	Han et al. (2018)	Relative error

General comparisons of DL models

This study identifies the most efficient DL model in each identified F&B domain. Table 4 illustrates our comparisons of the error terms in the pool of reviewed articles.

Note that "A > B" means that the performance of model A is better than that of model B. "A + B" indicates the hybridization of multiple DL models.

At this point, we have summarized three methods of data processing in DL models against seven specified F&B domains, including data preprocessing, data inputs, and evaluation rules. Apart from the technical level of DL, we find the following:

(1) NN has advantages in handling cross-sectional data;
(2) RNN and LSTM are more feasible in handling time series data;
(3) CNN has advantages in handling the data with multicollinearity.

Apart from application domains, we can induce the following viewpoints. Cross-sectional data usually appear in exchange rate prediction, price prediction, and macroeconomic prediction, for which NN could be the most feasible model. Time series data usually appear in stock market prediction, for which LSTM and RNN are the best options. Regarding stock trading, a feasible DL model requires the capabilities of decision and self-learning, for which RL can be the best. Moreover, CNN is more suitable for the multivariable environment of any F&B domains. As shown in the statistics of the Appendix, the frequency of using corresponding DL models corresponds to our analysis above. Selecting proper DL models according to the particular needs of financial analysis is usually challenging and crucial. This study provides several recommendations.

We summarize emerging DL models in F&B domains. Nevertheless, can these models refuse the efficient market hypothesis (EMH)?[5] According to the EMH, the financial market has its own discipline. There is no long-term technical tool that could outperform an efficient market. If so, using DL models may not be practical in long-term trading as it requires further experimental tests. However, why do most of the reviewed articles argue that their DL models of trading outperform the market returns? This argument has challenged the EMH. A possible explanation is that many DL algorithms are still challenging to apply in the real-world market. The DL models may raise trading opportunities to gain abnormal returns in the short-term. In the long run, however, many algorithms may lose their superiority, whereas EMH still works as more traders recognize the arbitrage gap offered by these DL models.

Discussion

This section discusses three aspects that could affect the outcomes of DL models in finance.

---

[5]EMH was developed from a Ph.D. dissertation by economist Eugene Fama in the 1960s. It says that at any given time, stock prices reflect all available information and trade at exactly their fair value at all times. It is impossible to consistently choose stocks that will beat the returns of the overall stock market. Therefore, this hypothesis implies that the pursuit of market-beating performance is more about chance than it is about researching and selecting the right stocks.

Table 4 Comparisons of the error terms based on our collected articles

Identified F&B domains	Comparisons and ranking of commonly-used DL models
Exchange rate Prediction	DBN > FNN
	DBN + RBM > RW, ARMA, FNN
	FNN + chaos theory + multi-objective evolutionary algorithms > FNN
	CNN > MA
Prediction of stock market	Generative algorithm + FNN > FNN, SVM
	RNN + Stock2vec > LSTM, FNN
	CNN > DT, SVM
	CNN + special order feature set > LR > CNN
	CNN > LSTM > RNN, ridge regression, Lasso, elastic net, random forecast, SVR, AdaBoost, gradient boosting, FNN, SVM, kNN, VAR, p-RBM.
	LSTM + GARCH-type model > GARCH, dxponential weighted MA, LSTM
	FNN + autoencoder + RBM > FNN, extreme learning machine, radial basis FNN
	FNN + PCA > RNN, radial basis FNN
Stock trading	FNN > buy-and-hold method
	FNN > ARMA-GARCH
	Fuzzy learning + RNN > CNN + RNN, LSTM + wavelet transforms + LSTM + SAEs > LSTM, RNN
	RL + Q-learning > buy-and-hold method
	RL + SARSA > buy-and-hold method
	GA + RNN+ RL > RL
	RL + DNN + transfer learning > RL, DNN
Portfolio management,	ListNet > RankNet
	RNN + RL > RNN
Banking default risk and credit	FNN + NLP > NLP
	CNN + feature selection > LR
	LSTM + Event2vec > random forest+ Event2vec
Macroeconomic prediction	FNN > DT, LR
	FNN > SVM, random forest, classification tree
	MLP (FNN) is the most widely used model in price prediction on oil.

Training and validation of data processing

The size of the training set

The optimal way to improve the performance of models is by enhancing the size of the training data. Bootstrap can be used for data resampling, and generative adversarial network (GAN) can extend the data features. However, both can recognize numerical parts of features. Sometimes, the sample set is not diverse enough; thus, it loses its representativeness. Expanding the data size could make the model more unstable. The current literature reported diversified sizes of training sets. The requirements of data size in the training stage could vary by different F&B tasks.

The number of input factors

Input variables are independent variables. Based on our review, multi-factor models normally perform better than single-factor models in the case that the additional input factors are effective. In the time-series data model, long-term data have less prediction errors than that for a short period. The number of input factors depends on the employment of the DL structure and the specific environment of F&B tasks.

The quality of data

Several methods can be used to improve the data quality, including data cleaning (e.g., dealing with missing data), data normalization (e.g., taking the logarithm, calculating the changes of variables, and calculating the t-value of variables), feature selection (e.g., Chi-square test), and dimensionality reduction (e.g., PCA). Financial DL models require that the input variables should be interpretable in economics. When inputting the data, researchers should clarify the effective variables and noise. Several financial features, such as technical indexes, are likely to be created and added into the model.

DL models

Selection on structures of DL models

DL model selection should depend on problem domains and cases in finance. NN is suitable for processing cross-sectional data. LSTM and other RNNs are optimal choices for time-series data in prediction tasks. CNN can settle the multicollinearity issue through data compression. Latent variable models like GAN can be better for dimension reduction and clustering. RL is applicable in the cases with judgments like portfolio management and trading. The return levels and outcomes on RL can be affected significantly by environment (observation) definitions, situation probability transfer matrix, and actions.

The setting of objective functions and the convexity of evaluation rules

Objective function selection affects training processes and expected outcomes. For predictions on stock price, low MAE merely reflects the effectiveness of applied models in training; however, it may fail in predicting future directions. Therefore, it is vital for additional evaluation rules for F&B. Moreover, it can be more convenient to resolve the objective functions if they are convex.

The influence of overfitting (underfitting)

Overfitting (underfitting) commonly happens in using DL models, which is clearly unfavorable. A generated model performs perfectly in one case but usually cannot replicate good performance with the same model and identical coefficients. To solve this problem, we have to trade off the bias against variances. Bias posits that researchers prefer to keep it small to illustrate the superiority of their models. Generally, a deeper (i.e., more layered) NN model or neurons can reduce errors.

However, it is more time-consuming and could reduce the feasibility of applied DL models.

One solution is to establish validation sets and testing sets for deciding the numbers of layers and neurons. After setting optimal coefficients in the validation set (Chong et al. 2017; Sevim et al. 2014), the result in the testing sets reveals the level of errors that could mitigate the effect of overfitting. One can input more samples of financial data to check the stability of the model's performance. This method is known as the early stopping. It stops training more layers in the network once the testing result has achieved an optimal level.

Moreover, regularization is another approach to conquer the overfitting. Chong et al. (2017) introduces a constant term for the objective function and eventually reduces the variates of the result. Dropout is also a simple method to address overfitting. It reduces the dimensions and layers of the network (Minh et al. 2017; Wang et al. 2019). Finally, the data cleaning process (Baek and Kim 2018; Bao et al. 2017), to an extent, could mitigate the impact of overfitting.

Financial models

The sustainability of the model

According to our reviews, the literature focus on evaluating the performance of historical data. However, crucial problems remain. Given that prediction is always complicated, the problem of how to justify the robustness of the used DL models in the future remains. More so, whether a DL model could survive in dynamic environments must be considered.

The following solutions could be considered. First, one can divide the data into two groups according to the time range; performance can subsequently be checked (e.g., using the data for the first 3 years to predict the performance of the fourth year). Second, the feature selection can be used in the data preprocessing, which could improve the sustainability of models in the long run. Third, stochastic data can be generated for each input variable by fixing them with a confidence interval, after which a simulation to examine the robustness of all possible future situations is conducted.

The popularity of the model

Whether a DL model is effective for trading is subject to the popularity of the model in the financial market. If traders in the same market conduct an identical model with limited information, they may run identical results and adopt the same trading strategy accordingly. Thus, they may lose money because their strategy could sell at a lower price after buying at a higher.

Conclusion and future works

Concluding remarks

This paper provides a comprehensive survey of the literature on the application of DL in F&B. We carefully review 40 articles refined from a collection of 150

articles published between 2014 and 2018. The review and refinement are based on a scientific selection of academic databases. This paper first recognizes seven core F&B domains and establish the relationships between the domains and their frequently-used DL models. We review the details of each article under our framework. Importantly, we analyze the optimal models toward particular domains and make recommendations according to the feasibility of various DL models. Thus, we summarize three important aspects, including data preprocessing, data inputs, and evaluation rules. We further analyze the unfavorable impacts of overfitting and sustainability when applying DL models and provide several possible solutions. This study contributes to the literature by presenting a valuable accumulation of knowledge on related studies and providing useful recommendations for financial analysts and researchers.

Future works

Future studies can be conducted from the DL technical and F&B application perspectives. Regarding the perspective of DL techniques, training DL model for F&B is usually time-consuming. However, effective training could greatly enhance accuracy by reducing errors. Most of the functions can be simulated with considerable weights in complicated networks. First, one of the future works should focus on data preprocessing, such as data cleaning, to reduce the negative effect of data noise in the subsequent stage of data training. Second, further studies on how to construct layers of networks in the DL model are required, particularly when considering a reduction of the unfavorable effects of overfitting and underfitting. According to our review, the comparisons between the discussed DL models do not hinge on an identical source of input data, which renders these comparisons useless. Third, more testing regarding F&B-oriented DL models would be beneficial.

In addition to the penetration of DL techniques in F&B fields, more structures of DL models should be explored. From the perspective of F&B applications, the following problems need further research to investigate desirable solutions. In the case of financial planning, can a DL algorithm transfer asset recommendations to clients according to risk preferences? In the case of corporate finance, how can a DL algorithm benefit capital structure management and, thus, maximize the values of corporations? How can managers utilize DL technical tools to gauge the investment environment and financial data? How can they use such tools to optimize cash balances and cash inflow and outflow? Until recently, DL models like RL and generative adversarial networks are rarely used. More investigations on constructing DL structures for F&B regarding preferences would be beneficial. Finally, the developments of professional F&B software and system platforms that implement DL techniques are highly desirable.

Appendix

Part A. Summary of publications in DL and F&B domains

Table 5 A summary of publications of DL models and F&B domains

	ANN/ DNN/ FNN/MLP	CNN	LSTM	RNN (Simple)	RL	DBN	RBM
Exchange rate prediction	Shen et al. (2015) Zheng et al. (2017) Ravi et al. (2017)	Galeshchuk and Mukherjee (2017)				Shen et al. (2015) Zheng et al. (2017)	
Stock market direction/ prediction/ movement/ volatility	Matsubara et al. (2018) Chen et al. (2018b) Chong et al. (2017) Yan and Ouyang (2017) Kim and Won (2018) Singh and Srivastava (2017)	Dingli and Fournier (2017) Gunduz et al. (2017) Tadaaki (2018)	Hernandez and Abad (2018) Sohangir et al. (2018) Krausa and Feuerriegel (2017) Yan and Ouyang (2017) Kim and Won (2018) Minh et al. (2017) Baek and Kim (2018)	Krausa and Feuerriegel (2017) Minh et al. (2017) Singh and Srivastava (2017) Jiang et al. (2018)			Chong et al. (2017) Hernandez and Abad (2018)
Stock trading	Sezer et al. (2017) Chen et al. (2017) Deng et al. (2017) Fischer and Krauss (2017) Hsu et al. (2018) Jeong and Kim (2018)	Deng et al. (2017)	Deng et al. (2017) Bao et al. (2017) Fischer and Krauss (2017)	Deng et al. (2017) Bao et al. (2017) Jeong and Kim (2018)	Deng et al. (2017) Martinez-Miranda et al. (2016) Feuerriegel and Prendinger (2016) Chakraborty (2019) Zhang and Maringer (2015) Chen et al. (2018a)		
Price prediction	Sehgal and Pandey (2015)						
Portfolio management	Song et al. (2017)			Almahdi and Yang (2017)	Almahdi and Yang (2017)		
Macroeconomic prediction	Sevim et al. (2014) Chatzis et al. (2018)						
Banking default risk and credit	Rönnqvist and Sarlin (2017)	Zhu et al. (2018)	Wang et al. (2019) Jurgovsky et al. (2018)		Han et al. (2018)		

Part B. Detailed structure of standard RNN

The abstract structure of RNN for a sequence cross over time can be extended, as shown in Fig. 7 in Appendix, which presents the inputs as X, the outputs as Y, the weights as w, and the Tanh functions.

Fig. 7 The detailed structure of RNN
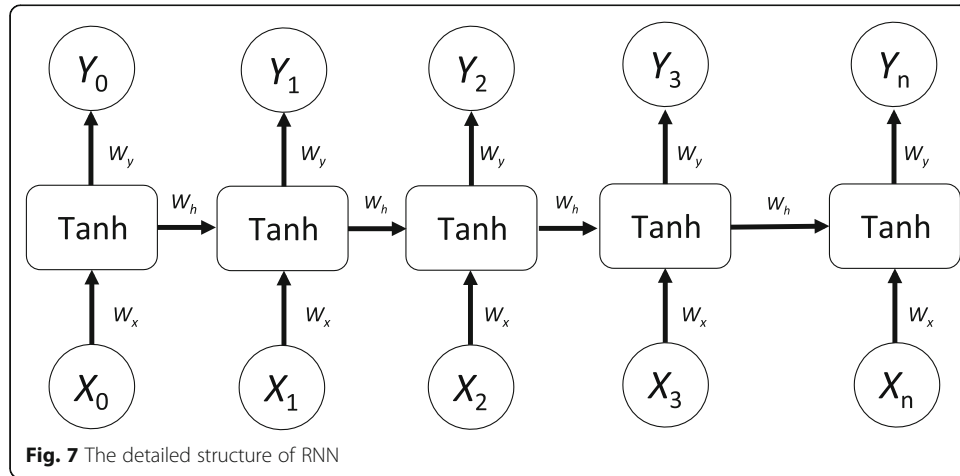
Part C. List of abbreviations

Table 6 A summary table of abbreviations used in this paper

ACC	Actual correlation coefficient	MAPE	Mean absolute percent error
ANN	Artificial neural network	MCC	Matthews correlation coefficient
ARMA	Auto-regressive-moving-average	ML	Machine learning
CNN	Convolutional neural networks	MLP	Multi-layer perceptron
DBN	Deep belief networks	MSE	Mean squared error
DL	Deep learning	NMSE	Normalized mean square error
DNN	Deep neural network	NN	Neural network
DPA	Directional predictive accuracy	PCA	Principal component analysis
EMH	Efficient market hypothesis	RBM	Restricted Boltzmann machine
FNN	Feedforward neural network	RL	Reinforcement learning
GA	Genetic algorithm	RMSE	Root mean squared error
GAN	generative adversarial network	RNN	Recurrent neural network
GARCH	Generalized autoregressive conditional heteroscedasticity	RNN(O)	Other recurrent neural networks except for LSTM
kNN	k-nearest neighbors	SAEs	Stacked autoencoders
LSTM	Long short-term memory	SVM	Support vector machine
MA	Moving average	VAR	Vector autoregression
MAE	Mean absolute error		

Authors' contributions
JH carried out the collections and analyses of the literature, participated in the design of this study and preliminarily drafted the manuscript. JC initiated the idea and research project, identified the research gap and motivations, carried out the collections and analyses of the literature, participated in the design of this study, helped to draft the manuscript and proofread the manuscript. SC participated in the design of the study and the analysis of the literature, helped to draft the manuscript and proofread the manuscript. The authors read and approved the final manuscript.

Availability of data and materials
Not applicable.

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

Author details
[1]Department of Mathematics, The Hong Kong University of Science and Technology, Hong Kong, China. [2]Division of Business and Management, BNU-HKBU United International College, Zhuhai, China.

References
Almahdi, S., & Yang, S. Y. (2017). An adaptive portfolio trading system: A risk-return portfolio optimization using recurrent reinforcement learning with expected maximum drawdown. *Expert Systems with Applications, 87*, 267–279.

Baek, Y., & Kim, H. Y. (2018). ModAugNet: A new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module. *Expert Systems with Applications, 113*, 457–480.

Bao, W., Yue, J., & Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short-term memory. *PLoS One, 12*(7), e0180944.

Butaru, F., Chen, Q., Clark, B., Das, S., Lo, A. W., & Siddique, A. (2016). Risk and risk management in the credit card industry. *Journal of Banking & Finance, 72*, 218–239.

Cavalcante, R. C., Brasileiro, R. C., Souza, V. L. F., Nobrega, J. P., & Oliveira, A. L. I. (2016). Computational intelligence and financial markets: A survey and future directions. *Expert System with Application, 55*, 194–211.

Chai, J. Y., & Li, A. M. (2019). Deep learning in natural language processing: A state-of-the-art survey. In *The proceeding of the 2019 international conference on machine learning and cybernetics* (pp. 535–540). Japan: Kobe.

Chai, J. Y., Liu, J. N. K., & Ngai, E. W. T. (2013). Application of decision-making techniques in supplier selection: A systematic review of literature. *Expert Systems with Applications, 40*(10), 3872–3885.

Chai, J. Y., & Ngai, E. W. T. (2020). Decision-making techniques in supplier selection: Recent accomplishments and what lies ahead. *Expert Systems with Applications, 140*, 112903. https://doi.org/10.1016/j.eswa.2019.112903.

Chakraborty, S. (2019). *Deep reinforcement learning in financial markets* Retrieved from https://arxiv.org/pdf/1907.04373.pdf. Accessed 04 Apr 2020.

Chatzis, S. P., Siakoulis, V., Petropoulos, A., Stavroulakis, E., & Vlachogiannakis, E. (2018). Forecasting stock market crisis events using deep and statistical machine learning techniques. *Expert Systems with Applications, 112*, 353–371.

Chen, C. T., Chen, A. P., & Huang, S. H. (2018a). Cloning strategies from trading records using agent-based reinforcement learning algorithm. In *The proceeding of IEEE international conference on agents* (pp. 34–37).

Chen, H., Xiao, K., Sun, J., & Wu, S. (2017). A double-layer neural network framework for high-frequency forecasting. *ACM Transactions on Management Information Systems, 7*(4), 11.

Chen, L., Qiao, Z., Wang, M., Wang, C., Du, R., & Stanley, H. E. (2018b). Which artificial intelligence algorithm better predicts the Chinese stock market? *IEEE Access, 6*, 48625–48633.

Chong, E., Han, C., & Park, F. C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems with Applications, 83*, 187–205.

Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., & Kuksa, P. (2011). Natural language processing (almost) from scratch. *Journal of Machine Learning Research, 12*, 2493–2537.

Deng, Y., Bao, F., Kong, Y., Ren, Z., & Dai, Q. (2017). Deep direct reinforcement learning for financial signal representation and trading. *IEEE Transactions on Neural Networks and Learning Systems, 28*(3), 653–664.

Dingli, A., & Fournier, K. S. (2017). Financial time series forecasting—A machine learning approach. *International Journal of Machine Learning and Computing, 4*, 11–27.

Elad, M., & Aharon, M. (2006). Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing, 15*(12), 3736–3745.

Feuerriegel, S., & Prendinger, H. (2016). News-based trading strategies. *Decision Support Systems, 90*, 65–74.

Fischer, T., & Krauss, C. (2017). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research, 270*(2), 654–669.

Galeshchuk, S., & Mukherjee, S. (2017). Deep networks for predicting the direction of change in foreign exchange rates. *Intelligent Systems in Accounting, Finance and Maangement, 24*(4), 100–110.

Gunduz, H., Yaslan, Y., & Cataltepe, Z. (2017). Intraday prediction of Borsa Istanbul using convolutional neural networks and feature correlations. *Knowledge-Based Systems, 137*, 138–148.

Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., & Lew, M. S. (2016). Deep learning for visual understanding: A review. *Neurocomputing, 187*, 27–48.

Han, J., Jentzen, A., & Weinan, E. (2018). Solving high-dimensional partial differential equations using deep learning. *The proceedings of the National Academy of Sciences of the United States of America (PNAS)*; 8505–10).

Hernandez, J., & Abad, A. G. (2018). Learning from multivariate discrete sequential data using a restricted Boltzmann machine model. In *The proceeding of IEEE 1st Colombian conference on applications in computational intelligence (ColCACI)* (pp. 1–6).

Hsu, P. Y., Chou, C., Huang, S. H., & Chen, A. P. (2018). A market making quotation strategy based on dual deep learning agents for option pricing and bid-ask spread estimation. *The proceeding of IEEE international conference on agents* (pp. 99–104).

Jeong, G., & Kim, H. Y. (2018). Improving financial trading decisions using deep Q-learning: Predicting the number of shares, action strategies and transfer learning. *Expert Systems with Applications, 117*, 125–138.

Jiang, X., Pan, S., Jiang, J., & Long, G. (2018). Cross-domain deep learning approach for multiple financial market predictions. *The proceeding of international joint conference on neural networks* (pp. 1–8).

Jurgovsky, J., Granitzer, M., Ziegler, K., Calabretto, S., Portier, P. E., Guelton, L. H., & Caelen, O. (2018). Sequence classification for credit-card fraud detection. *Expert Systems with Applications, 100*, 234–245.

Kim, H. Y., & Won, C. H. (2018). Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. *Expert Systems with Applications, 103*, 25–37.

Krausa, M., & Feuerriegel, S. (2017). *Decision support from financial disclosures with deep neural networks and transfer learning* Retrieved from https://arxiv.org/pdf/1710.03954.pdf Accessed 04 Apr 2020.

Krauss, C., Do, X. A., & Huck, N. (2017). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P500. *European Journal of Operational Research, 259*(2), 689–702.

Martinez-Miranda, E., McBurney, P., & Howard, M. J. W. (2016). Learning unfair trading: A market manipulation analysis from the reinforcement learning perspective. In *The proceeding of 2016 IEEE conference on evolving and adaptive intelligent systems (EAIS)* (pp. 103–109).

Matsubara, T., Akita, R., & Uehara, K. (2018). Stock price prediction by deep neural generative model of news articles. *IEICE Transactions on Information and Systems, 4*, 901–908.

Minh, D. L., Sadeghi-Niaraki, A., Huy, H. D., Min, K., & Moon, H. (2017). Deep learning approach for short-term stock trends prediction based on two-stream gated recurrent unit network. *IEEE Access, 6*, 55392–55404.

Ravi, V., Pradeepkumar, D., & Deb, K. (2017). Financial time series prediction using hybrids of chaos theory, multi-layer perceptron and multi-objective evolutionary algorithms. *Swarm and Evolutionary Computation, 36*, 136–149.

Rönnqvist, S., & Sarlin, P. (2017). Bank distress in the news describing events through deep learning. *Neurocomputing, 264*(15), 57–70.

Sehgal, N., & Pandey, K. K. (2015). Artificial intelligence methods for oil price forecasting: A review and evaluation. *Energy System, 6*, 479–506.

Sevim, C., Oztekin, A., Bali, O., Gumus, S., & Guresen, E. (2014). Developing an early warning system to predict currency crises. *European Journal of Operational Research, 237*(3), 1095–1104.

Sezer, O. B., Ozbayoglu, M., & Gogdu, E. (2017). A deep neural-network-based stock trading system based on evolutionary optimized technical analysis parameters. *Procedia Computer Science, 114*, 473–480.

Shen, F., Chao, J., & Zhao, J. (2015). Forecasting exchange rate using deep belief networks and conjugate gradient method. *Neurocomputing, 167*, 243–253.

Singh, R., & Srivastava, S. (2017). Stock prediction using deep learning. *Multimedia Tools Application, 76*, 18569–18584.

Sohangir, S., Wang, D., Pomeranets, A., & Khoshgoftaar, T. M. (2018). Big data: Deep learning for financial sentiment analysis. *Journal of Big Data, 5*(3), 1–25.

Song, Q., Liu, A., & Yang, S. Y. (2017). Stock portfolio selection using learning-to-rank algorithms with news sentiment. *Neurocomputing, 264*, 20–28.

Tadaaki, H. (2018). Bankruptcy prediction using imaged financial ratios and convolutional neural networks. *Expert Systems with Applications, 117*, 287–299.

Wang, C., Han, D., Liu, Q., & Luo, S. (2019). A deep learning approach for credit scoring of peer-to-peer lending using attention mechanism LSTM. *IEEE Access, 7*, 2161–2167.

Yan, H., & Ouyang, H. (2017). Financial time series prediction based on deep learning. *Wireless Personal Communications, 102*, 683–700.

Zhang, J., & Maringer, D. (2015). Using a genetic algorithm to improve recurrent reinforcement learning for equity trading. *Computational Economics, 47*, 551–567.

Zheng, J., Fu, X., & Zhang, G. (2017). Research on exchange rate forecasting based on a deep belief network. *Neural Computing and Application, 31*, 573–582.

Zhu, B., Yang, W., Wang, H., & Yuan, Y. (2018). A hybrid deep learning model for consumer credit scoring. In *The proceeding of international conference on artificial intelligence and big data* (pp. 205–208).

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.