

Retail Sales Forecasting Using Deep Learning: Systematic Literature Review

Linda Eglite and Ilze Birzniece*

Institute of Applied Computer Systems, Riga Technical University,
Meza 1 k1, Riga, LV-1048, Latvia

Linda.Eglite_1@edu.rtu.lv, Ilze.Birzniece@rtu.lv

Abstract. This systematic literature review examines the deep learning (DL) models for retail sales forecast. The accuracy of a retail sales forecast is a prevalent force for uninterrupted business operations. Accuracy for retailers means limiting supply chain and storage costs, ensuring no product is out of stock, and facilitating smooth promotional operations. The study analyses the DL frameworks used in reviewed literature. Tested DL models are listed, as well as other machine learning and linear models used for the evaluation comparison. Additionally, the review presents the metrics used by the authors for the model evaluation. This article concludes by describing the benefits and limitations of DL models for sales forecasting.

Keywords: Deep Learning, Sales Forecasting, Retail.

1 Introduction

To stay competitive retail companies must look for ways of increasing the efficiency of operations. For any retailer the main focus of the business is the volumes of sales, consequently having precise sales estimates is the cornerstone of every business. Without it, the supply chain, finance, marketing, or any other function in the company cannot operate without disruptions. As a result of underestimated sales, the product may end out of stock, marketing activities can be disrupted, and customers can be lost; in turn, overestimation can cause problems with the shelf life of products, and ultimately increase the cost of storage, products, and operations. Thus, sales forecast accuracy trickles down to the overall efficiency of the business. With ever-growing technological capabilities, it makes the most sense for retailers to look for solutions in this area.

Artificial intelligence (AI), although it can now be found almost everywhere – in our phones, laptops, cars, watches – still has many unexplored and insufficiently developed applications. One of the most advanced technologies of AI is deep learning (DL). DL as technology has a variety of applications such as image recognition, speech recognition, natural language understanding, acoustic modeling, and prediction modeling [1].

* Corresponding author

© 2022 Linda Eglite and Ilze Birzniece. This is an open access article licensed under the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>).

Reference: L. Eglite and I. Birzniece, “Retail Sales Forecasting Using Deep Learning: Systematic Literature Review,” *Complex Systems Informatics and Modeling Quarterly*, CSIMQ, no. 30, pp. 53–62, 2022. Available: <https://doi.org/10.7250/csimq.2022-30.03>

Additional information. Author’s ORCID iD: I. Birzniece – <https://orcid.org/0000-0002-5775-6138>. PII S225599222200170X. Received: 7 February 2022. Accepted: 8 April 2022. Available online: 30 April 2022.

This article is a piece of secondary research with the aim of identifying, evaluating, and interpreting currently available research regarding the usage of DL.

The goal of this systematic literature review is to summarize the existing knowledge regarding retail sales forecasting using DL technology and to provide an evaluation of benefits and limitations of the approaches used in DL for retail sales forecasting.

To achieve this goal the following research questions have been identified:

- RQ1. What are the DL models used for sales forecasting?
- RQ2. What metrics are used for model evaluation?
- RQ3. What are the benefits of using DL models for sales forecasting?
- RQ4. What are the challenges and limitations of using DL models for sales forecasting?

The structure of the article is as follows. Section 2 introduces sales forecasting, the approaches used for estimating it, with a specific focus on DL concepts. Section 3 presents the research method and search strategy. Section 4 shows the process of literature search and article selection. Section 5 contains the data analysis and results, and Section 6 concludes the article.

2 Background

This section provides the background to sales forecasting and the approaches used, additionally giving a brief overview of DL models and metrics commonly used.

According to Mentzer & Moon [2], a sales forecast is a “projection into the future of expected demand, given a stated set of environmental conditions”. In some of the earlier works, instead of “sales forecast”, the terms “sales prediction” or “demand forecast” have been used as synonyms of “sales forecasting”. The general approach of using time series historical data for estimating the future value of sales is the common factor in the reviewed literature, therefore, for this study, “sales forecasting” as an umbrella term will be applied.

For addressing forecasting problems, different models and methods have been used. Two of the classical forecasting methods are Auto-Regressive Integrated Moving Average (ARIMA), Seasonal Auto-Regressive Integrated Moving Average (SARIMA), where exponential smoothing performs statistical time series analysis. These are often used for market-level sales forecasts [3], [4].

Besides the conventional methods, there are also methods based on machine learning (ML). ML is a subset of AI. ML algorithms rely on data and learning to reach a specific goal by extracting patterns from the data. Some of the ML models are Linear Regression, k-Nearest Neighbor (k-NN), Random Forest (RF), and Support Vector Machine (SVM) [5], [6]. Artificial neural networks (ANN) are a specific discipline within ML. DL is an even smaller part of AI and ML. DL focuses on multilayer ANN and uses them as a backbone for DL algorithms [5]. Since 2006 the third popularity wave of ANN algorithms has started, and the term “deep learning” solidified its presence in the academic literature [7]. Some of the DL architectures commonly used are Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and Recurrent Neural Network (RNN). Schmidhuber [7] provides a comprehensive historical overview of the development of DL in NN. Additionally, more information on the DL models and architectures can be retrieved from [5] and [6]. Figure 1 describes the inclusive nature of the AI-related terms.

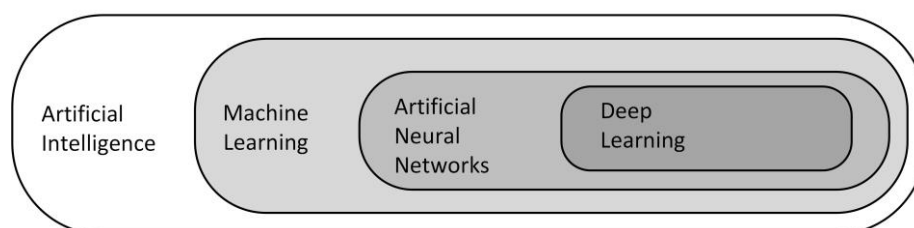


Figure 1. Relationships between AI components

The ultimate goal of classification is prediction. Varied measures are used to capture the quality of the prediction or forecast. Mean Absolute Error (MAE) calculates the average size of the error that the forecast contains (see Formula 1). Additionally, the Mean Absolute Percentage Error (MAPE) shows the average absolute percent of how far a forecast value is off the actual sales (Formula 2). As it is expressed as a percentage, MAPE allows evaluation of the overall accuracy of the model and comparison with other models. One of the typical measures used to evaluate the error of a model in predicting numerical data is Root Mean Square Error (RMSE), which enables the comparison of the predicted value with the actual observation for different models (Formula 3). A lower RMSE value means that the model has been able to forecast the values within a smaller error range and thus fits the data the best. The formulas (1–3) describe the most common metrics where a_t is the actual sales value, while f_t is the value forecasted by the model.

$$MAE = \frac{1}{n} \sum_{i=1}^n |a_t - f_t| \quad (1)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|a_t - f_t|}{a_t} \times 100 \quad (2)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (a_t - f_t)^2} \quad (3)$$

Additional information on accuracy measures is available in [8].

Emmert-Streib et al. [1] summarized the DL architectures used for every kind of prediction model. Their article focuses on the theory of DL and the various ways in which it can be applied. Also, Fildes et al., [4] wrote about the methods and aspects of retail forecasting not, however, focusing on DL. To the best of our knowledge, no work so far has analyzed the literature regarding retail sales forecasting with the use of DL prediction models, and moreover, there is a lack of comprehensive analysis of DL models and their applicability in this field.

3 Research Method

The systematic literature review process is performed according to Kitchenham and Charters report [9]. Figure 2 displays the main phases of the method.

The process starts with the research need identification. During this phase, the focus of the study has been selected, and relevance established. The process continues with specifying research questions. This step is important as it sets the framework of the research scope and findings. The next phase is to develop a search strategy. This serves as a roadmap for the research to find all the relevant literature and show the completeness, rigor, and transparency of the process. As shown in Figure 2, the next task is to perform the literature analysis of which the main part is data extraction and amalgamation. Further, we move to presenting study results followed by a discussion which answers the research questions set in Section 1. Lastly, we derive conclusions from the study.

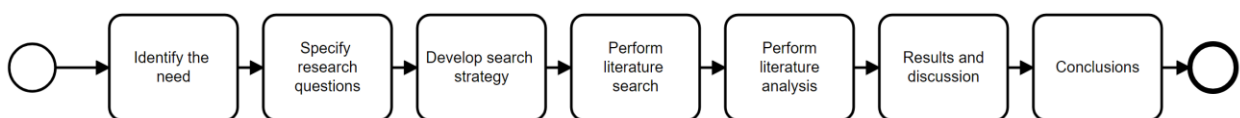


Figure 2. Main research phases according to [9]

3.1 Search Strategy

The following search strategy is used for the selection of relevant studies. First, a list of keywords is compiled. The list consists of word groups derived from research questions, synonyms of the words, a preliminary review of the topic in Scopus database, and the taxonomy of IEEE.

Identified keywords: deep learning, deep neural network, retail, purchasing prediction model, sales prediction, predictive models, prediction modeling, prediction methods, sales forecasting.

Second, these keywords are used for search string development. The search string was used to search article titles, keywords, and abstracts.

Search string: (Deep learning OR Deep Neural network) AND Retail AND (Purchasing prediction model OR sales prediction OR Predictive models OR Prediction modeling OR prediction methods OR sales forecasting)

Third, the search string was used in the following digital libraries: Scopus, IEEE Explorer, ACM Digital Library, and Science Direct.

Fourth, **the exclusion and inclusion criteria**, to which the studies had to comply were stated :

- Criterion 1** Articles in the selected databases must have free access provided by Riga Technical University or be open/public, otherwise, the article is excluded.
- Criterion 2** Articles must be written in English. Articles written in other languages are excluded.
- Criterion 3** An article must be written after the year 2005, articles prior this year are excluded. The year is chosen as it gives a sufficient period for review and, around that time, the term “deep learning” started to gain popularity [1].
- Criterion 4** An article must be a conference proceeding or journal article, other types of works like books, standards, and courses are excluded.
- Criterion 5** An article must be relevant to the topic and subject area of retail sales prediction using DL.

4 Literature Search and Selection

Figure 3 presents the literature search strategy implementation process in the selected databases and the criteria applied to the studies. Initially, using the search string, 137 studies were identified, further, with the criteria of language, publication year, and source type, 100 studies were selected for further examination. These 100 studies were examined based on title, abstract, and full access rights. In total 19 articles have been selected after applying the search strategy inclusion and exclusion criteria described in the previous section.

Criterion	Action	# of studies
Criterion 1	Search in selected databases Use of search string	137
Criterion 2	Filter Language English	136
Criterion 3	Filter Year 2005 +	121
Criterion 4	Source type Conference proceeding or journal article	100
Criterion 5	Abstract and access Abstract relevant to subject, full access	19

Figure 3. Literature search process

5 Findings and Results

The reviewed literature concerned data from different retail businesses and industries. Table 1 displays the industries represented and the number of studies reviewed from those industries. Grocery, e-commerce, apparel and accessory industry, alcohol, health and beauty, and shopping mall sales have been forecasted by the authors. Most frequently authors used Python for the model implementation. Overall, the 19 studies reviewed are dating from 2016 to 2021, so it is the most current academic literature that has been considered. The following sections will introduce the analysis of the extracted data, answer the research questions, and present the results.

Table 1. Industry prevalence in the reviewed literature

Industry	Number of articles	Reference
Alcohol retail	1	[10]
Apparel & Accessories	2	[11], [12]
E-commerce	3	[13]–[15]
Grocery stores	9	[3], [16]–[23]
Health & Beauty	2	[24]
Not specified retail	2	[25], [26]
Shopping Mall	1	[27]

5.1 Prediction Models

DL can be achieved by using many different neural network architectures. As a subset of machine learning, DL uses perceptron, heuristics, and it often utilizes large datasets. The most common architectures of DL are ANN, RNN, CNN, LSTM, and MLP. Each one is slightly different in the tasks it can perform and its architectural complexity [1].

The studies of prediction modeling used various techniques and developed additional frameworks based on the above-mentioned architectures. One study used the K-means algorithm for data clustering and LSTM architecture for the prediction model. This combination allowed the model to reach a high level of prediction accuracy even with limited historical data [5]. Kaneko and Yada [16] use a simple DL framework to predict whether the sales would increase or decrease.

Table 2 consists of the DL models considered by reviewed articles, and machine learning or linear models used for the comparison. DL models are divided into two parts. The first 11 entries listed and marked grey in Table 2 are original frameworks or architecture adaptation proposals from the reviewed papers' authors (H2O, DSF, STANet, ASFC, NN MPL, EE-CNN, EMD-G, EMD-MG, NN Model, AGA-LSTM, CNN-LSTM). The following 7 models have established DL architectures that are used to perform the analysis in a novel setting or make an extensive model comparison with DL, ML, and linear models, like [11], [13]. The most common DL model used by the authors is LSTM architecture. It is a type of RNN architecture, which is applicable to many uses including natural language processing, voice recognition, and, of course, predictions [1].

Table 2. List of models used in reviewed studies

		MODELS, FRAMEWORKS																																				
		DEEP LEARNING											OTHER MACHINE LEARNING					LINEAR																				
		H2O framework	DSF	STANet	ASFC	NN MLP	EE-CNN	EMD-G	EMD-MG	NN Model	AGA-LSTM	CNN-LSTM	LSTM	DNN	MLP	ANN	RNN	CNN(CNN+WD)	CatBoost	Linear Regression	SVM	RF	Logistic Regression	XGBoost	Decision trees	kNN	LSVR	Ridge	SVR	GBDT	GBM	GBRT	ARIMA	AR	ARMA	EMA	Exponential smoothing	
Articles	[16]	X																				X															2	
	[13]		X											X				X													X	X	X		X		7	
	[22]											X					X				X											X					4	
	[26]			X																							X	X					X				4	
	[3]				X													X													X		X				4	
	[17]												X								X	X															3	
	[23]															X																						1
	[11]													X		X					X	X		X				X									6	
	[12]					X																															1	
	[25]						X														X	X	X														4	
	[15]												X																								1	
	[24]													X																	X						2	
	[27]							X	X							X						X	X	X					X								7	
	[18]																	X	X	X																	3	
	[19]									X											X	X															3	
	[20]												X	X							X	X	X								X				X	7		
	[14]										X	X																									2	
[21]											X	X					X																			3		
[10]												X								X	X															3		
		1	1	1	1	1	1	1	1	1	1	7	3	2	2	2	2	2	1	6	6	4	2	2	1	1	1	1	1	1	1	1	4	2	1	1	1	

AGA-LSTM – Adaptive Genetic Algorithm LSMT; ANN – Artificial Neural Networks; AR –Autoregressive; ARIMA – Auto-Regressive Integrated Moving Average; ARMA – Auto-Regressive Moving Average; ASFC – Aggregate sales forecasting framework based on ConvLSTM; CNN+WD – Convolutional Neural Network with weights assigned; DNN – Deep Neural Network; DSF – Deep Sales Forecasting Framework; EE-CNN – Entity Embedding and convolutional neural network; EMA – Exponential Moving Average; EMD-G – Empirical Mode Decomposition and GRU method; EMD-MG – Empirical Mode Decomposition and MLP (multi-Layer perceptron) and GRU method; GBDT – Gradient Boosting Decision Tree; GBM – Gradient Boosting Method; GBRT – Gradient boosting regression tree; kNN – k-Nearest Neighbor; LSTM – Long Short-Term Memory; LSVR – Linear Support Vector Regression; NN Model – Neural Network Model; NN MLP – Neural Network Multi-Layer perceptron; RF – Random Forest; RNN – Recurrent Neural Network; STANet – Structural Temporal Attention Network; SVM – Support Vector Machine; SVR – Support Vector Regression; XGBoost – eXtreme Gradient Boosting

5.2 Model Evaluation Metrics

In the reviewed articles it is common to compare the target DL models with some other baseline models. For this evaluation the authors use different types of evaluation metrics. This section analyses the metrics used to compare the prediction models. The summary of evaluation metrics used in the related work is displayed in Table 3.

Table 3. List of metrics used in reviewed studies

	Metrics																		
	RMSE	MAE (or MAD)	Accuracy*	MAPE (wMApe)	MSE	AUC	F-Measure	Precision	Recall	Generalization ability	Runtime	Cost	Convenience	RSE	RMSSE	R ²	EMP	Model training time	
Articles	[16]			x		x	x	x	x									5	
	[11]	x	x		x	x										x		5	
	[22]			x						x	x	x	x					5	
	[12]	x	x			x												3	
	[3]	x	x		x													3	
	[23]		x		x	x												3	
	[25]			x			x										x	3	
	[20]	x																x	2
	[27]	x	x																2
	[13]		x		x														2
	[24]	x	x																2
	[19]	x																	1
	[14]					x													1
	[18]	x																	1
	[21]			x															1
	[17]														x				1
	[26]													x					1
[15]	x																	1	
[10]			x															1	
	9	7	5	4	4	2	1	1	1	1	1	1	1	1	1	1	1	1	

AUC – Area Under the Curve; EMP – Expected Maximum Profit; MAD – Mean Absolute Deviation; MAE – Mean Absolute Error; MAPE – Mean Absolute Percentage Error; MSE – Mean Squared Error; RMSE – Root Mean Squared Error; RMSSE – Scaled Root Mean Squared Error; RSE – Relative Squared Error; wMAPE – Weighted Mean Absolute Percentage Error
 Accuracy* - no particular accuracy measure is denoted in the article and cannot be identified

Figure 4 displays, as a percentage, how many articles used a particular metric. The most common metric used is RMSE; this metric has been used by almost half of the surveyed articles (47%, 9 articles). MAE has been used by 37% or 7 studies. The articles usually refer to 2-3 metrics in their analysis. Most commonly the authors use a combination of RMSE, MAE, MAPE or MSE to achieve the most accurate evaluation.

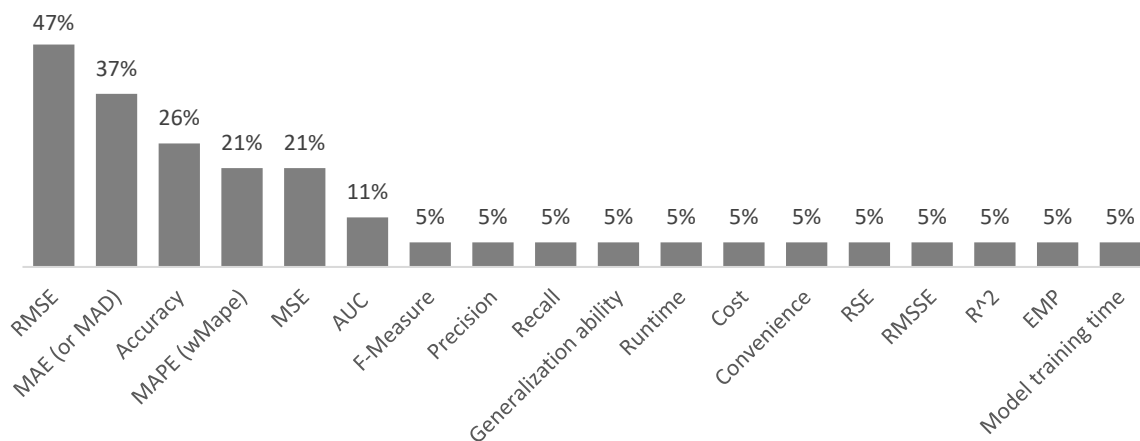


Figure 4. Frequency of metrics

5.3 DL Benefits

This section addresses the benefits that are achieved by applying deep learning to the prediction model. DL prediction models open new capabilities, that might not be possible to reach with standard models, for instance, Giri et al. [14] proposed a neural network MPL model, that by examining apparel features in the product picture can estimate the sales of new, previously unsold products. STANet framework developed by Liao et al. [26] achieves superiority in sales prediction accuracy by considering, not only historical sales, but also relationships among products. They hypothesized that a relationship among products, like being of the same brand, would play a role in the sales forecast. Qi et al. [13] and Chena et al. [19] similarly consider product relationships and promotions in their model to achieve higher accuracy. DL models allow the estimation of data with non-linear relationships, which is one of the biggest advantages over linear models [19].

However, the most important benefit is apparent from the comparative analysis of DL and other approaches. As displayed in Figure 5, 82% (14 articles) of the research articles that had DL model comparison with machine learning or linear models found that DL had superior results. Two papers found that other types of model had a better result, while one study did not have a conclusive answer, as the results varied based on metrics.

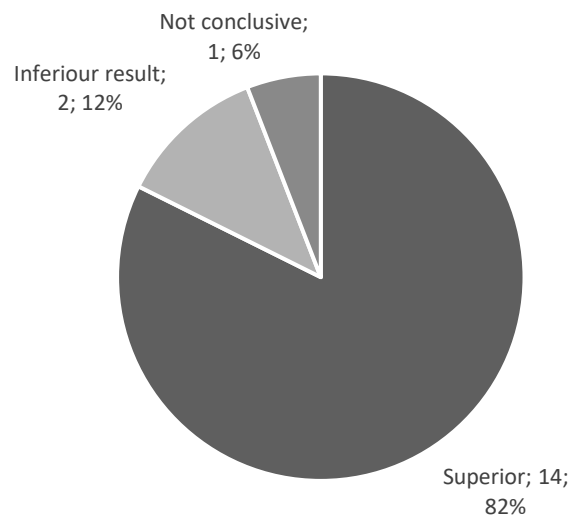


Figure 5. DL model result vs comparison models

5.4 DL Limitations

This section considers the challenges and limitations of using DL models for sales predictive modeling. Some specific limitations can be mentioned for each of the reviewed approaches. For instance, the Kaneko and Yada [16] model was able to predict increase or decrease, but not the actual sales figure. However, this also means that their model was simple and easy to implement. For more complex problems, the DL models do become extensively complex to implement and understand. DL models are the so-called black-box solution; thus, their decision-making process is not traceable, and the approach would not be applicable if knowledge extraction and outcome explanation is required [19]. Additionally, similarly to non-neural network ML models, DL models may also be subject to biases, overfitting (the model being trained too close to given data) or underfitting (model not suitable for generalization and not being able to model the data) [5]. That is the reason why so many DL algorithms are still being developed, evaluated, and tested. Thus, although the potential of DL models is there, the limitations cannot be ignored.

6 Conclusions

This study is a summary of sales forecast models using DL. Based on the article reviews performed in this study, it is possible to see that varied models and evaluation metrics have been applied to the grocery store, e-commerce, apparel and accessory, health and beauty store, shopping mall, and alcohol sales forecasts. The DL architectures, used most commonly for sales forecast, are LSTM, DNN, and MPL. However, other authors choose to develop their own frameworks (11 cases were found). For grocery store sales forecasting, in 4 out of 9 applications, LSTM was chosen, while 4 out of 9 applications used novel frameworks. For the evaluation of the models, the authors of reviewed articles have compared DL applications with non-neural network ML algorithms and linear models. Grocery store forecast comparison is most commonly done with the linear Regression model (4 instances), SVM (4 instances), and ARIMA (3 instances). For other industries, the data pool is too small to draw any tendencies regarding the most prevalent models for sales forecasting. The most often used metrics for evaluation are RMSE, MAE, and MAPE; however other metrics are accepted and used as well. The most frequently used combination of metrics is RMSE and MAE (5 instances). In some of the studies, the particular accuracy measure used was not denoted and could not be identified thus pointing to the lack of consistent and repeatable research methodology of these studies. Application of DL frameworks proves to provide superior sales forecast methods and allows for capabilities not possible with other methods; however, most often, they are complex solutions and are harder to implement than other ML or linear models. Additional research for improving the models and broadening the application areas is needed and it is clear that DL architectures are still developing, and we can expect to see more works using DL in the near future. This research can serve as a basis for further development of retail sales prediction models, taking into consideration the findings, strengths, and weaknesses of up-to-date solutions using DL, other ML, and linear models.

References

- [1] F. Emmert-Streib, Z. Yang, H. Feng, S. Tripathi, and M. Dehmer, "An Introductory Review of Deep Learning for Prediction Models with Big Data," *Frontiers in Artificial Intelligence*, vol. 3, p. 4, 2020. Available: <https://doi.org/10.3389/frai.2020.00004>
- [2] J. T. Mentzer Jr. and M. A. Moon, *Sales Forecasting Management: A Demand Management Approach*. Thousand Oaks, United States: Sage Publications, 2004. Available: <https://doi.org/10.4135/9781452204444>
- [3] Y. Liu, K. Lan, F. Huang, X. Cao, B. Feng, and B. Zhu, "An Aggregate Store Sales Forecasting Framework based on ConvLSTM," *ACM International Conference Proceeding Series*, pp. 67–72, 2021. Available: <https://doi.org/10.1145/3456529.3456540>
- [4] R. Fildes, S. Ma, and S. Kolassa, "Retail forecasting: Research and practice," *International Journal of Forecasting*, 2019. Available: <https://doi.org/10.1016/j.ijforecast.2019.06.004>
- [5] O. Campesato, *Artificial intelligence, machine learning, and deep learning*. Mercury Learning & Information, 2020.
- [6] I. J. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. Available: <http://www.deeplearningbook.org>. Accessed on Feb. 04, 2022.
- [7] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85–117, Jan. 2015. Available: <https://doi.org/10.1016/j.neunet.2014.09.003>
- [8] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International Journal of Forecasting*, vol. 22, no. 4, pp. 679–688, Oct. 2006. Available: <https://doi.org/10.1016/j.ijforecast.2006.03.001>
- [9] B. Kitchenham and S. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering," 2007.
- [10] A. Palkar, M. Deshpande, S. Kalekar, and S. Jaswal, "Demand Forecasting in Retail Industry for Liquor Consumption using LSTM," *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, pp. 521–525, 2020. Available: <https://doi.org/10.1109/ICESC48915.2020.9155712>

- [11] A. L. D. Loureiro, V. L. Miguéis, and L. F. M. da Silva, “Exploring the use of deep neural networks for sales forecasting in fashion retail,” *Decision Support Systems*, vol. 114, pp. 81–93, 2018. Available: <https://doi.org/10.1016/j.dss.2018.08.010>
- [12] C. Giri, S. Thomasse, J. Balkow, and X. Zeng, “Forecasting New Apparel Sales Using Deep Learning and Nonlinear Neural Network Regression,” *2019 International Conference on Engineering, Science, and Industrial Applications (ICESI)*, pp. 1–6, 2019. Available: <https://doi.org/10.1109/ICESI.2019.8863024>
- [13] Y. Qi, C. Li, H. Deng, M. Cai, Y. Qi, and Y. Deng, “A deep neural framework for sales forecasting in e-commerce,” *International Conference on Information and Knowledge Management, Proceedings*, pp. 299–308, 2019. Available: <https://doi.org/10.1145/3357384.3357883>
- [14] K. Chen, “An Online Retail Prediction Model Based on AGA-LSTM Neural Network,” *2020 2nd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI)*, pp. 145–149, 2020. Available: <https://doi.org/10.1109/MLBDBI51377.2020.00032>
- [15] Y. Lingxian, K. Jiaqing, and W. Shihuai, “Online retail sales prediction with integrated framework of K-mean and neural network,” *PervasiveHealth: Pervasive Computing Technologies for Healthcare*, pp. 115–118, 2019. Available: <https://doi.org/10.1145/3345035.3345048>
- [16] Y. Kaneko and K. Yada, “A Deep Learning Approach for the Prediction of Retail Store Sales,” *IEEE International Conference on Data Mining Workshops, ICDMW*, pp. 531–537, 2016. Available: <https://doi.org/10.1109/ICDMW.2016.0082>
- [17] X. Li, J. Du, Y. Wang, and Y. Cao, “Automatic Sales Forecasting System Based on LSTM Network,” *Proceedings – 2020 International Conference on Computer Science and Management Technology, ICCSMT 2020*, pp. 393–396, 2020. Available: <https://doi.org/10.1109/ICCSMT51754.2020.00088>
- [18] J. Ding, Z. Chen, L. Xiaolong, and B. Lai, “Sales Forecasting Based on CatBoost,” *Proceedings - 2020 2nd International Conference on Information Technology and Computer Application, ITCA 2020*, pp. 636–639, 2020. Available: <https://doi.org/10.1109/ITCA52113.2020.00138>
- [19] J. Chen, W. Koju, S. Xu, and Z. Liu, “Sales Forecasting Using Deep Neural Network and SHAP techniques,” *2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering, ICBAIE 2021*, pp. 135–138, 2021. Available: <https://doi.org/10.1109/ICBAIE52039.2021.9389930>
- [20] Z. Huo, “Sales Prediction based on Machine Learning,” *Proceedings – 2nd International Conference on E-Commerce and Internet Technology, ECIT 2021*, pp. 410–415, 2021. Available: <https://doi.org/10.1109/ECIT52743.2021.00093>
- [21] P. Kaunchi, T. Jadhav, Y. Dandawate, and P. Marathe, “Future Sales Prediction for Indian Products Using Convolutional Neural Network-Long Short Term Memory,” *2021 2nd Global Conference for Advancement in Technology (GCAT)*, pp. 1–5, 2021. Available: <https://doi.org/10.1109/GCAT52182.2021.9587668>
- [22] J. Wang, G. Q. Liu, and L. Liu, “A Selection of Advanced Technologies for Demand Forecasting in the Retail Industry,” in *2019 4th IEEE International Conference on Big Data Analytics, ICBDA 2019*, pp. 317–320, 2019. Available: <https://doi.org/10.1109/ICBDA.2019.8713196>
- [23] R. C. Syafitri, A. Y. Ridwan, and N. Novitasari, “Demand Forecasting for Drinking Water Products to Reduce Gap between Estimation and Realization of Demand Using Artificial Neural Network (ANN) Methods in PT. XYZ,” *Proceedings of the International Conference on Engineering and Information Technology for Sustainable Industry (ICONETSI)*, pp. 1–6, 2020. Available: <https://doi.org/10.1145/3429789.3429844>
- [24] K. Wanchoo, “Retail Demand Forecasting: A Comparison between Deep Neural Network and Gradient Boosting Method for Univariate Time Series,” *2019 IEEE 5th International Conference for Convergence in Technology (I2CT)*, pp. 1–5, 2019. Available: <https://doi.org/10.1109/I2CT45611.2019.9033651>
- [25] B. Zhu, W. Tang, X. Mao, and W. Yang, “Location-based Hybrid Deep Learning Model for Purchase Prediction,” *Proceedings – 2020 5th International Conference on Computational Intelligence and Applications, ICCIA 2020*, pp. 161–165, 2020. Available: <https://doi.org/10.1109/ICCIA49625.2020.00038>
- [26] S. Liao, X. Zhou, and W. Rao, “Accurate demand forecasting for retails with deep neural networks,” *Advances in Database Technology – EDBT2020*, pp. 455–458, 2020. Available: <https://doi.org/10.5441/002/edbt.2020.56>
- [27] S. Mou, Y. Ji, and C. Tian, “Retail Time Series Prediction Based on EMD and Deep Learning,” *Proceedings of 2018 6th IEEE International Conference on Network Infrastructure and Digital Content, IC-NIDC 2018*, pp. 425–430, 2018. Available: <https://doi.org/10.1109/ICNIDC.2018.8525707>