FMDB *Transactions on Sustainable Computer Letters*

Analysis of Machine Learning and Deep Learning Methods for Superstore Sales Prediction

S. Silvia Priscila[1,*], S. Suman Rajest[2], Sai Nitisha Tadiboina[3], R. Regin[4], Szeberényi András[5]

[1]Department of Computer Science, Bharath Institute of Higher Education and Research, Tamil Nadu, India.
[2]Department of Research and Development (R&D), Dhaanish Ahmed College of Engineering, Chennai, Tamil Nadu, India.
[3]Department of Information Technology, Geico, Maryland, USA.
[4]Department of Computer Science and Engineering, SRM Institute of Science and Technology, Chennai, Tamil Nadu, India.
[5]Institute of Marketing, Budapest Metropolitan University, Budapest, Hungary.
silviaprisila.cbcs.cs@bharathuniv.ac.in[1], sumanrajest414@gmail.com[2], stadiboina@geico.com[3], regin12006@yahoo.co.in[4], aszeberenyi@metropolitan.hu[5]

Abstract: Today, a group of supermarkets requires a consistent ridge of their yearly sales. This primarily results from a need for knowledge, resources, and the capability to estimate sales. Conventional statistical methods for supermarket sales are important and often lead to predictive models. In the age of big data and powerful computers, machine learning is the standard for sales forecasting. This comprehensive literature review examines superstore sales prediction models using ML and DL. This article review focuses on superstore sales prediction using machine learning and deep learning in data mining. Finally, DL is the best SSP for results. DL models market movements well. Automatic feature extraction models and forecasting strategies have been tested with various inputs. DL algorithms process large real-time datasets better. DL research found the best hybrid processing methods for real-time stock market data. DL and ML methods predict the client's response and identify its factors. DL and ML algorithms are evaluated using Rodolfo Saladanha marketing campaign data. Four metrics precision, recall, F-measure, and accuracy compare ML and DL algorithms. MATLAB tested these methods. LSTM, CNN, LR, RF, and LR algorithms were used to compare results to well-known ML and DL algorithms. Artificial Convolutional Neural Network (ACNN) is compared to RF, LR, CNN, and LSTM. The proposed superstore sales prediction algorithm outperformed the others. The proposed model predicted superstore sales with a validation accuracy of 93.90 percent, outperforming current and suitable baselines.

Keywords: Deep Learning (DL); Superstore Sales Prediction (SSP); Random Forest (RF); Linear Regression (LR); Convolutional Neural Network (CNN); Long Short-Term Memory Networks (LSTM); Deep Neural Network (DNN).

Received on: 06/10/2022, **Revised on:** 22/11/2022, **Accepted on:** 03/01/2023, **Published on:** 05/02/2023

Cited by: S. S. Priscila, S.S. Rajest, S. N. Tadiboina, R. Regin and S. András, "Analysis of Machine Learning and Deep Learning Methods for Superstore Sales Prediction," *FMDB Transactions on Sustainable Computer Letters*., vol. 1, no. 1, pp. 1–11, 2023.

1. Introduction

Sales forecasting is necessary for any successful retail firm in today's aggressive business environment, where sales are of the greatest significance to businesses. Retail businesses must seek ways towards progress working effectiveness if they want to stay aggressive. Since sales volumes are any retailer's primary concern, having precise sales predictions is essential to the success of every organization. By preventing overproduction and decreasing overstock, it can help in inventory management [1]. Also, it can be a tremendous tool designed for boosting a business' profitability during cost forecasts [2]. Several significant

[*]Corresponding author.

factors impact future sales. It is important to note that the level of predicting difficulty varies depending on the product. Milk products have reliable demand during the year, making it easy to forecast their sales [3]. Since sales volumes are the primary focus of any store, having accurate sales predictions is essential to the success of every enterprise.

Profit is the ultimate objective of every supermarket. This is accomplished when more products are sold, and there is a high turnover. Using this data, gold distinguishes high-performing supermarkets from the competition. The volume of sales data has been constantly growing in recent years. The data with the purpose of customers collecting when they engage with different supermarkets. The company's supply chain, finances, marketing, and other departments cannot run smoothly without it. Overestimation can result in issues with product shelf life, ultimately increasing the cost of storage, goods, and operations. Underestimation can result in the product running out of stock, marketing initiatives being disrupted, and customers being lost. So, the accuracy of sales forecasts has a ripple effect on business productivity. A key component of retail is shopping sales. Administrators rigorously evaluate the most trustworthy models to aid in forecasting future sales. Forecasting shop sales can help anticipate future sales changes or increases when making decisions based on historical and present data. Understanding sales patterns and trends will help businesses or merchants make more money and improve customer satisfaction through precise prediction [4].

The most reasonable course of action for merchants in this regard is to look for answers, given the rapidly expanding capabilities of technology. The forecasting of supermarket sales can be done in several ways, but historically, many supermarkets have relied on these conventional statistical models. However, Machine Learning (ML) has developed into a significant field of data science that has gained popularity because of its strong predictive and forecasting abilities. As a result, it has become the go-to for extremely precise sales forecasting and other crucial fields [5]. Machine Learning (ML), created in the information science sector recently, is used in practical contexts like data mining. Researchers continue to glean knowledge and insights that are helpful to businesses from the significant amounts of consumer information data that these businesses have amassed utilizing these data mining techniques, coupled with big data, an area of study. Because they are implemented without understanding customers' purchasing behaviours, traditional sales and marketing goals do not help firms keep up with the pace of a competitive market. In the studies mentioned above, there are primarily three issues that stand out:

- It is challenging to segment customers based on their relationship with particular products;
- Effective feature representation from raw data; and
- All of the features were manually chosen and required expert knowledge.

Yet, there has been interest in using DL techniques for this issue due to the increasing processing power offered by current computers. For the purpose of analyzing Supermarket Marketing Campaigns with high reliability and classification accuracy, DL approaches are presented. Applications for DL as a technology include predictive modeling, acoustic modeling, image recognition, speech recognition, and natural language comprehension [6].

The developments in DL have greatly impacted the worlds of sales and marketing. Superstore sales prediction using machine learning and superstore sales prediction using deep learning in data mining are the two main headings of this article review. Lastly, the ideal SSP for effective results is the DL methodology. The remainder of this essay is broken out as follows. The literature review provided in the following section 2 addresses various ML and DL algorithms for sales prediction. The effectiveness of various approaches is compared in Section 3 and results and discussion in 4. Finally, the conclusion and future work are covered in Section 5.

2. Review of Machine Learning Methods For SSP

Odegua [7] introduced using K Nearest Neighbor (KNN), Gradient Boosting (GB), and Random Forest (RF) for estimating grocery sales. The introduced effort consists of two main steps: classification, data processing, and engineering. One-hot encoding has been introduced to encode all categorical variables, the mean was used to fill in any missing instances in the Product Weight feature, and the data set was then standardized by removing the mean from the data and dividing the result by the standard deviation. Second, using three machine learning algorithms-KNN, GB, and RF, forecasted sales for the grocery chain "Chukwudi Supermarkets." Supermarket kind (Grocery Store), product pricing, supermarket opening year, and supermarket location, among others, were the most crucial factors that aided in better sales forecasting. The data set comprises a sample of 4990 occurrences with 13 features/variables. Performance evaluation metrics are Standard Deviation (SD) and Mean Absolute Error (MAE). Regression is the type of work, and the MAE is a tried-and-true statistic that accurately assesses model performance. The outcomes demonstrate that the RF algorithm outperforms the GB and KNN models.

Tom et al. [8] created an ensemble classifier using RF, Decision Trees (DT), Ridge Regression (RR), XGboost, and Linear Regression (LR) to predict product sales. The connection between the imputed qualities is reduced when missing values are replaced with means and modes because they are numerical. Second, regardless of the sort of prediction job, whether it is regression or classification, an ensemble classifier employing RF, DT, RR, XGboost, and LR is implemented. Many ML algorithms have used 2013 Big Mart Sales data to forecast sales in the future. This model's accuracy analysis is contrasted with those of other models.

Singh et al. [9] introduced a MapReduce (MR) to analyze Walmart's sales data. To ascertain the business drivers, use the Walmart Shop. It can forecast which departments will be impacted by certain circumstances and how they will affect sales at various stores. Large Walmart sales datasets kept in Comma Separated Value (CSV) format are among the tools and methods used for this endeavour. Apache Spark uses Hadoop Distributed File Systems (HDFS) with a build of Hadoop as a data storage option. Apache Spark is a framework that can manage batch and stream processing simultaneously on the same application. Instead of using the standard Spark shell, InteliJ Idea was integrated with Spark. A wider audience can now easily analyze huge amounts of data thanks to the spark data frame Application Programming Interface (API), which also supports distributed data processing in general-purpose programming languages like Scala, Python, and Java. A spark-Structured Query Language (SQL) component has been introduced to query the data because the files are stored in data frames. Lastly, use Spark-SQL to apply MR functions on the datasets. Throughput and response time are the performance indicators.

Wang [10] formulated an Adaptive-Neuro-Fuzzy Inference System (ANFIS) for predicting consumer electronics sales. These nonlinear outputs are handled by ANFIS, which also provides sustainable development and management. This system provides decision-making that takes several objectives and multiple outputs into account. Analyzing several variables, including the design process, client specifications, computational intelligence, and soft computing, focuses on real-world problem-solving. The ANFIS inference module calculates the threshold value using all the inputs from related sensors. The outputs from sensor nodes are connected to the server to configure consumer electronics. The sensors may record variables like temperature, frequency, motion, sound, video, light, pressure, etc. The performance evaluation metrics are Threshold Value and Response Time.

Cheriyan et al. [11] proposed a Generalized Linear Model (GLM), Decision Tree (DT), and Gradient Boost Tree (GBT) for predicting sales. Data collection, preparation, and prediction are the introduced work steps. Data preparation and collecting came first. The initial dataset included a lot of entries, but after removing useless, redundant, and irrelevant data, the final dataset was significantly smaller than the initial dataset. Second, prediction is concerned with future events. Three ML algorithms, GLM, DT, and GBT, can be used for prediction. The three years of sales data from an e-fashion shop were used to create the dataset for this study. Past sales data for the three years from 2015 to 2017 were gathered for the e-fashion store to anticipate future sales. The performance evaluation metrics are accuracy rate, error rate, precision, recall and Kappa. According to the studies, GBT is the best-fit model, showing the highest accuracy and lowest error rate in forecasting and future sales prediction.

Ching-Chin et al. [12], resolve the new product sales forecasting problem, there are two methods: the New Product Sales Forecasting Process (NPSFP) and the New Product Forecasting System (NPFS). The NPSFP approach standardizes the sales forecasting processes by directing data collection and analysis, selecting the forecasting model, calculating real forecasts, and manually adjusting forecasting results. NPFS outperformed the Moving Average approach. Results are measured using Mean Absolute Percentage Error (MAPE).

Choi et al. [13] introduced a Grey Model (GM) and Extended Extreme Learning Machine (EELM) for retail fashion. The EELM and GM processes are combined in an intelligent forecasting algorithm. The introduced real data study shows that this newly created method may produce accurate predictions given the available time and data constraints. Due to the trend factor's importance and the seasonal cycle's typical high variability in quick fashion, these two characteristics fit the fast fashion demand pattern quite well. The sales data was gathered from a knitwear fashion business that adopted the rapid fashion idea. Results are measured using Mean Absolute Percentage Error (MAPE). The findings may benefit Future fast fashion operations from real-time sales forecasting. There are also some managerial ramifications mentioned.

Zuo et al. [14] formulated a Support Vector Machine (SVM) and Naive Bayes (NB) classifiers that are used to predict what customers will buy in a grocery store. In the proposed work, two representative machine learning techniques, NB classifier and SVM, are used to examine how well they perform with real-world data. SVM is a member of the group of theories known as supervised learning, which is comparatively very effective for tasks involving classification, regression, and clustering. An SVM classifier's distinctive kernel component sets it apart from other learning algorithms in that it can efficiently handle high-dimensional data space. Results are measured using accuracy.

Yang and Ramingwon [15] proposed an RF regression for supermarket sales forecasting. Pearson Correlation Coefficient (PCC) is introduced to calculate the relationship between features and then rank the source data with important information of random forest features to get the most important features to the model, i.e., the influencing factors. The original data used in this study came from publicly accessible data on Alibaba's Tianchi platform, which included sales information from a small shop with three branches. The decidability factor, Mean Squared Error (MSE), Mean Absolute Error (MAE), and Explained Variance Score (EVS) were selected as indicators for the model evaluation.

Wang [16] developed an XGboost, and LightGBM created a sales forecasting model. Preprocessing, feature selection, and classification are the introduced work steps. Initially, preprocessing is required since the dataset includes non-numeric data that must be transformed into numeric values before it can be used with the model. The second step is the introduction of Recursive Feature Elimination (RFE), which uses cross-validation to establish the ideal number of features. Then carried out model integration and gave weights based on the predictions made by these two models. Results are measured using Root Mean Squared Error (RMSE). According to experimental findings, this method's RMSE is 2.07, much lower than the RMSE of the two models used before the integration. The RMSE of the model based on LightGBM is 2.09, and the RMSE based on the XGBoost framework is 2.11.

Ali et al. [17] introduced RF, Decision Tree (DT), and Linear Regression (LR) models that are used to forecast big-mart sales. It entails the following actions: generating hypotheses, data exploration, data cleansing, and feature engineering are the first three steps. By examining the sales of other markets with particular characteristics, you may forecast the market's sales. Concatenate files to better comprehend the data because the dataset consists of two distinct train and test files. According to the experimental findings, of the four machine learning algorithms used, the RF predictor performed better than the RR, LR, and DT models. The proposed models' effectiveness has been assessed using Root Mean Square Error (RMSE).

Jiménez et al. [18] introduced Online sales forecasting using the Evolutionary Non-dominated Radial slots-based Algorithm (ENORA), RF, Non-dominated Sorted Genetic Algorithm (NSGA-II), and RFE. Preprocessing, feature selection, and decision-making are the two main steps of the proposed study. Then, the modes and means from the training data replace all the missing values for nominal and numerical variables. Second, the ENORA wrapper technique uses feature selection as a search approach. Lastly, the RF model introduces a precise regression model. Test and evaluate ENORA's performance as a multi-objective evolutionary search strategy compared to the original dataset, a traditional backward search strategy like RFE, and other multi-objective evolutionary search strategies like NSGA-II. The Kaggle repository is where the dataset was gathered. RMSE has evaluated the performance of the proposed models.

Pane et al. [19] formulated an association rule in supermarket sales that can be improved using Particle Swarm Optimization with Gradient Boosting Regression (PSO-GBR) based on a regression model that can forecast weekly sales. Compare the introduced GBR with the two established prediction techniques, Linear Regression (LR) and AdaBoost Regression (ABR), as a benchmark for this research. Preparing the data comes first. Finally, using GBR, create a model to forecast sales using the dataset. The model will next be put through a benchmark evaluation. Next, use PSO to maximize sales using an optimal regression technique. The final stage demonstrates that the strategy yields the best results for Weekly Sales. The supermarket dataset was produced using Kaggle's "Retail Data Analytics." GBR has a higher R2 score for the train and test datasets, respectively, of 0.95 and 0.94. Next, it is demonstrated that PSO increases sales using the GBR model's cost function.

3. Review of Deep Learning Methods for SPP

Pan and Zhou [20] proposed E-commerce sales forecasting using Convolutional Neural Networks (CNN). To forecast commodities sales, CNN is used to mine e-commerce data. The first step in this process is to integrate the intrinsic properties of the pertinent product data with the original cargo log dataset, which can be transformed into a particular "data frame" format. The raw log file includes items sold over a long period, price, quantity view, browsing, searching, times collected, the number of products added to the cart, and many other metrics. The data frame is then processed using CNN to extract useful characteristics. CNN's final layer employs these attributes to forecast product sales. The 1814,892 records in the experimental dataset that the Alibaba Group has made available. On a genuine e-commerce dataset, the proposed algorithm's viability is confirmed. Mean Square Error (MSE) is a performance evaluation statistic used in experiments.

Kaneko and Yada [21] developed an L1 regularisation deep learning, a recent development in the fast-expanding field of machine learning that has attracted much interest. A retail Store's Point-of-Sale (POS) data collected over three years was utilized to create a sales prediction model that, given a day's sales, forecasts a change in sales for the following day. Consequently, a deep learning model that considers L1 regularisation had the highest accuracy rate for predicting sales, which was 86.00%. Three years' worth of POS data from supermarkets in the Kanto region of Japan was gathered between 2002 and 2004. Results are evaluated by the precision, recall, F-measure, accuracy, and Area Under of Curve (AUC).

Samonte et al. [22] introduced an LSTM, and RF is used to forecast sales in retail establishments. Data preprocessing, data analysis, and data testing are the introduced work steps. First, preprocess the data by removing any rows with incorrect or missing values and dividing them into training and test sets. Subsequently, data analysis proposed a system that, when dealing with time-series data to estimate sales, predominantly used the LSTM algorithms. Also, the proposed system used the RF algorithms as a baseline for comparison with the applied prediction model in the study. In order to increase the predictive accuracy, the introduced system focused on forecasting sales in retail establishments that used four-stage procedures to generate the dataset. The Kaggle dataset was obtained and altered to fit the features. RMSE and MAPE measures are used to assess and contrast the performance of the two models, LSTM and Autoregressive Integrated Moving Average (ARIMA).

Punam et al. [23] formulated a two-level strategy to estimate product sales from a certain outlet. The proposed labour processes, which include feature engineering, data transformation, and (1) data exploration, are crucial for correctly forecasting outcomes. During the data exploration phase, it was discovered that the values for Item Weight and Outlet Size were missing. The weight of that specific item was averaged to fill in any missing values for item weight, and the mode of the outlet size for that specific type of outlet was used to fill in any missing values for outlet size. Big Mart Sales data from 2013 is used as the dataset for the technique. As an evaluation statistic, MAE has been used to examine the models' performance.

Leo Alexander and Delwin Christopher [24] proposed Big Mart sales forecasting using Linear Regression (LR), Principle Component Analysis (PCA), RF, SVM, and Neural Network (NN). The introduced work steps include feature engineering and data cleaning. Following data cleaning, missing values are imputed using multivariate chained equations (MICE). The ensemble-based prediction model was then created by feature engineering using fundamental statistical predictive models like the generally linear, PCA-based, and other machine learning techniques like RF, SVM, and NN. The forecast is generated using one data set, for which the historical sales are available, while the analysis uses a different data set. The information was gathered through Analytics Vidhya, and RMSE was utilized to analyze the outcomes.

Zhaoweijie et al. [25] developed A significant supermarket chain XGboost gradient model for Rossmann. The sales position of Rossmann Company's 1115 stores over the last three years predicts how much money will be spent in those 1115 stores over the following 48 days. The two main steps of the proposed study are feature engineering and data pre-processing. Before creating the spatial dimension features, pre-processing, data transformation, and association establish the time dimension characteristics. Lastly, performs feature engineering from the perspective of space and time, builds a prediction model using the XGboost gradient promotion technique, and predicts the sales. The model's accuracy in predicting the future sales of 41088 samples in the Kaggle competition was 89.07%.

Niu [26] introduced an XGboost, and careful feature engineering processing is combined in the XGboost sale prediction model, which foresees Walmart's sales issues. To efficiently make accurate forecasts, use the properties of many dimensions. Walmart supermarket sales data has been collected from the Kaggle competition datasets to evaluate the XGboost sale forecast model. Experimental findings demonstrate that the introduced methods outperform alternative ML strategies. Route Mean Square Scale Error (RMSSE) in the proposed study is 0.141 and 0.113 less than that of the LR and ridge algorithms.

Jha and Pande [27] formulated an increase in sales in a commercial setting; Facebook (FB) Prophet and Auto-Regressive Integrated Moving Average (ARIMA) are used. For predicting sales, the regression model and the logistic exponential model has been introduced in this work. The most recent technique to demonstrate enhanced supermarket forecast accuracy is FB Prophet. The proposed models' performance accuracy is evaluated using three matrices, MSE, RMSE, and MAPE. It may be inferred from the introduced research that the ARIMA system is a better prediction model due to its reduced error, better prediction, and better fitting.

Kaunchi et al. [28] proposed a Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) hybrid for sales prediction. A real-time big mart dataset from the neighbourhood market stores is used to test the proposed model. The estimations collected introduced that the number created matches the actual data, with a maximum degree of precision of 97% and a minimum degree of precision exceeding 82.00%, according to the predictive approach.

Dai and Huang [29] proposed an LSTM for predicting sales. Employ an LSTM model with a unique loss function and the hyper-parameter search for accuracy optimization. The performance of two different LSTMs with or without the hyper-parameter search is compared using six different machine learning models using Auto Machine Learning (AutoML) as the baseline. Use the open dataset, Kaggle Rossman sales data, to demonstrate the performance. Based on user preferences that reflect the needs of the actual business, it can also legitimately exaggerate or underestimate sales estimates. Prediction accuracy, the Root Mean Square Percentage Error (RMSPE), and the MAPE are used to analyze the results.

Chen [30] developed an Adaptive Genetic Algorithm with Long Short-Time Memory (AGA-LSTM) neural network-based online retail prediction model. In order to increase the model's predictive accuracy, the AGA is used during the LSTM construction process to optimize network parameters like the time step, the number of hidden layers, and the training times. The model then forecasts the different types of goods and the overall sales volume. Results from the University of California, Irvine's Online Retail II dataset have been gathered (UCI). It keeps track of every online transaction order data from UK shops from December 1, 2010, to December 9, 2011. The results demonstrate that the AGA-LSTM model's prediction accuracy is significantly higher than that of the conventional LSTM model, demonstrating the efficiency of this method by MSE.

Lingxian et al. [31] introduced an integrated framework for online retail sales prediction that combines conventional unsupervised learning of data features with K-mean clustering and a cutting-edge artificial intelligence model of the LSTM. Lastly, note that by utilizing a modest quantity of historical data, the modelling approach can represent near-term online retail dynamics. The dataset includes the transaction records for an online retail operation in the UK between December 1, 2010, and December 9, 2011. Stochastic Gradient Descent (SGD) optimization and MSE are both employed.

Ding et al. [32] formulated a CatBoosting-based method for sales forecasting. The feature engineering and prediction system are the two main steps of the proposed study. A machine learning pipeline always starts with feature engineering. A decision tree gradient boosting approach with great performance is called CatBoosting. It is a quick and precise framework with built-in support for categorical features for gradient boosting on oblivious trees. The M5 Forecasting-Accuracy dataset from Kaggle was used to train the model. The dataset includes sales information for 30490 Walmart products during 1914 days. The introduced model outperforms established machine learning techniques like LR and SVM in the experiments, achieving an RMSE of 0. 605. The proposed method requires less fine-tuning than existing methods, which enhances its capacity to generalize to other custom datasets and increases its potential applications. The RMSE is the measurement tool for this competition.

Chen et al. [33] proposed a Walmart sales forecasting model using a Deep Neural Network (DNN). Moreover, test the NN model using the Kaggle platform's datasets. The network's neurons require input signals to be received and processed and output signals to be sent. Studies have revealed that the introduced NN model delivers greater performance by RMSE compared to other machine learning models. The RMSE metric outperforms the LR and SVM by 2.92 and 2.58, respectively. Additionally, the SHapley Additive exPlanations (SHAP) method is employed to analyze the NN model to mine features of various dimensions and create accurate predictions.

4. Results and Discussion

The RF, LR, CNN, LSTM, and ACNN classification algorithms have all been considered for experimentation. A confusion matrix, like the one in Table 1, can represent the details of the real and expected samples. Experiments employing the introduced deep learning classifier have employed several learning paradigms. The classification results show that the choice of characteristics and the classification model are crucial in predicting superstore sales. A superstore is making preparations for its holiday sale. Rodolfo Saladanha's marketing campaign dataset was used for classification.

Several performance metrics have been considered to validate the introduced model's performance, including precision, recall, F1-score (or) F-measure, accuracy, and confusion matrix as evaluation matrices. The confusion matrix listed in Table 1 determines the metric utilized for the two-class classification problem. A confusion matrix is a table with two rows and two columns used in predictive analytics to represent the number of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) outcomes of a classification model. TP and TN show the number of correct classifications. Nevertheless, FP and FN indicate the quantity of inaccurate classification.

Table 1: Representation of Confusion Matrix

	Predicted positive	**Predicted negative**
Actual positive	True Positive (TP)	False Negative (FN)
Actual negative	False Positive (FP)	True Negative (TN)

4.1. Dataset

A superstore is making preparations for its holiday sale. They want to introduce a new deal: a gold membership that costs $499 but is otherwise $999 and offers a 20% discount on all purchases. Only current customers will be eligible, and a phone campaign is being developed for them. The management believes that creating a predictive model that will categorise clients who might take advantage of the offer is the best method to lower the cost of the campaign. The superstore wants to determine the various elements influencing the customer's response and forecast the possibility that the consumer will respond favourably. The dataset was obtained from the Rodolfo Saldanha marketing campaign on Kaggle [34].

4.2. Evaluation Metrics

Precision, recall, accuracy, F-measure, and confusion matrix have been taken into account as performance metrics to convey the comparative results of ML and DL algorithms and validate the performance of the given model. These measurements have been calculated and applied using MATLAB. True positive and true negative in this study refer to the accurate forecast of a possible client and a non-potential consumer, respectively. False positive refers to a client who is not potential but is classified as such by the prediction model, and false negative refers to a customer who is potential but is classified as such by the prediction model.

Precision is defined as equation (1),

$$\text{Precision} = \frac{TP}{(TP + FP)} \tag{1}$$

Recall is defined as equation (2),

$$\text{Recall} = \frac{TP}{(TP + FN)} \tag{2}$$

F-measure is the harmonic mean of precision, and recall is defined as equation (3),

$$F1 = \frac{2 * \text{Precision} * \text{Recall}}{(\text{Precision} + \text{Recall})} \tag{3}$$

Accuracy is defined as equation (4),

$$\text{Accuracy} = \frac{\text{TP+TN}}{(\text{TP+TN+FP+FN})} \tag{4}$$

Here, TP and TN serve as examples of the quantity of accurate classification. FP and FN, however, indicate the number of incorrect classifications.

4.3. Results Comparison

Table 2 clearly outlines the performance comparison of the precision, recall, accuracy, F1-score, or F-measure values. The accompanying figures 1–4 plot these four metrics as simulation outcomes. Figures 1–4 provide evidence that the introduced methodology, ACNN, is evaluated against other current techniques, including RF, LR, CNN, and LSTM.

Table 2: Prediction Algorithms by Various Evaluation Metrics

Prediction Algorithms	Evaluation Metrics			
	Precision (%)	**Recall (%)**	**F-measure (%)**	**Accuracy (%)**
LR	69.55	71.18	70.63	74.94
SVM	74.67	75.49	75.08	77.57
CNN	77.38	80.72	79.05	80.48
LSTM	81.54	83.29	82.41	84.62
ACNN	83.95	87.84	85.89	88.45

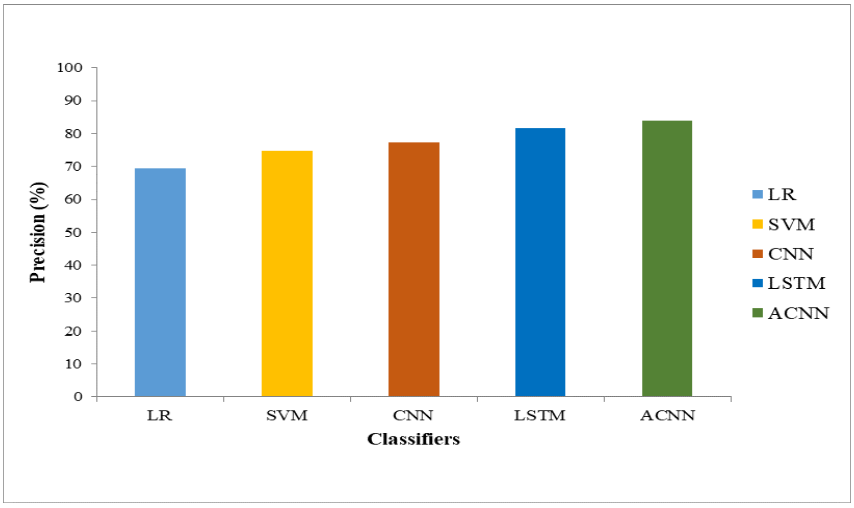

Figure 1: Precision Performance Comparison in Various Classification Methods

Figure 1 displays the precision comparison between the introduced ACNN and existing approaches in various classification methods, including RF, LR, CNN, and LSTM. Compared to other approaches already in use, the proposed method can achieve a high precision rate, as shown in Figure 1. It is the best method for predicting supermarket sales, with an 83.9% accuracy rate. Compared to the introduced method, various existing methods, such as RF, LR, CNN, and LSTM, provide precision rates of 66.5%, 69.6%, 78.3%, and 81.5%, respectively.

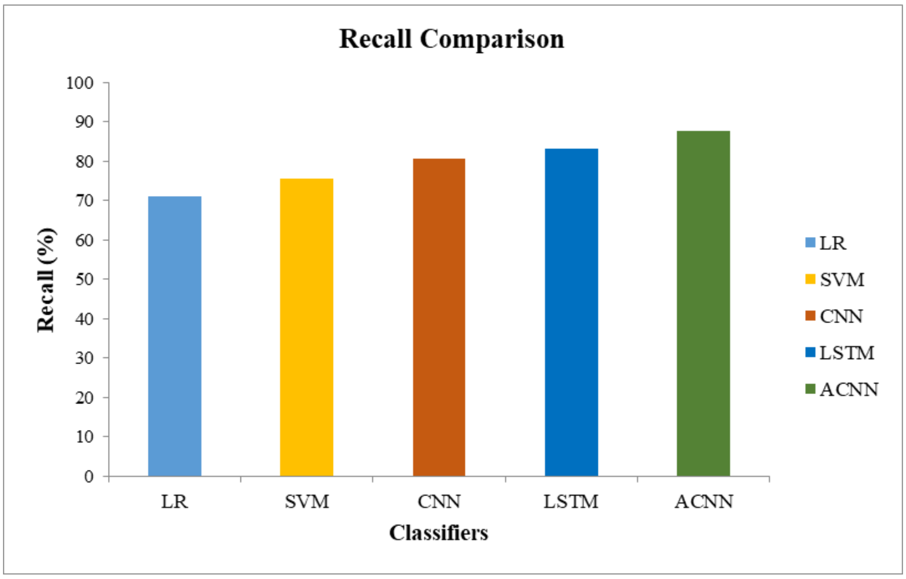

Figure 2: Recall Performance Comparison in Various Classification Methods

Figure 2 illustrates the recall comparison in several classification approaches between the proposed ACNN and current methods, including RF, LR, CNN, and LSTM. Compared to other current methods, the introduced method can achieve a high recall rate, as shown in picture 2. With a high recall rate of 87.8%, it is the best method for predicting grocery store sales. Compared to the proposed method, the recall rates provided by the existing approaches, such as RF, LR, CNN, and LSTM, are, respectively, 70.5%, 80.5%, 82.7%, and 84.1%.

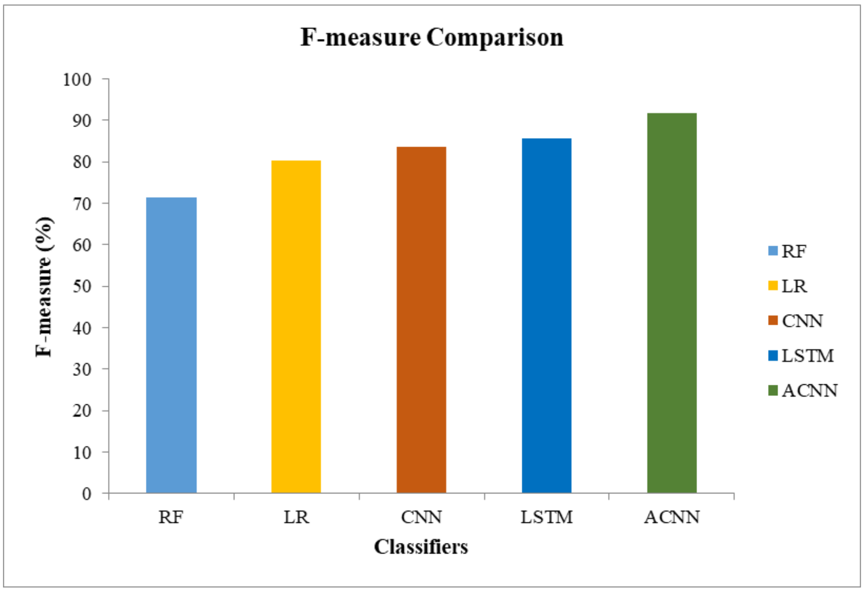

Figure 3: F-Measure Performance Comparison in Various Classification Methods

Figure 3 illustrates the F-measure comparison in several classification methods between the proposed ACNN and current methods, including RF, LR, CNN, and LSTM. Compared to other current approaches, the introduced method can achieve a high F-measure rate, as shown in Figure 3. It is the most accurate method of predicting supermarket sales, with a high F-measure rate of 91.8%. The F-measure rates provided by the ACNN compared to the RF, LR, CNN, and LSTM are 71.3%, 80.4%, 83.5%, and 85.7%, respectively; nevertheless, these rates are lower than the introduced technique.

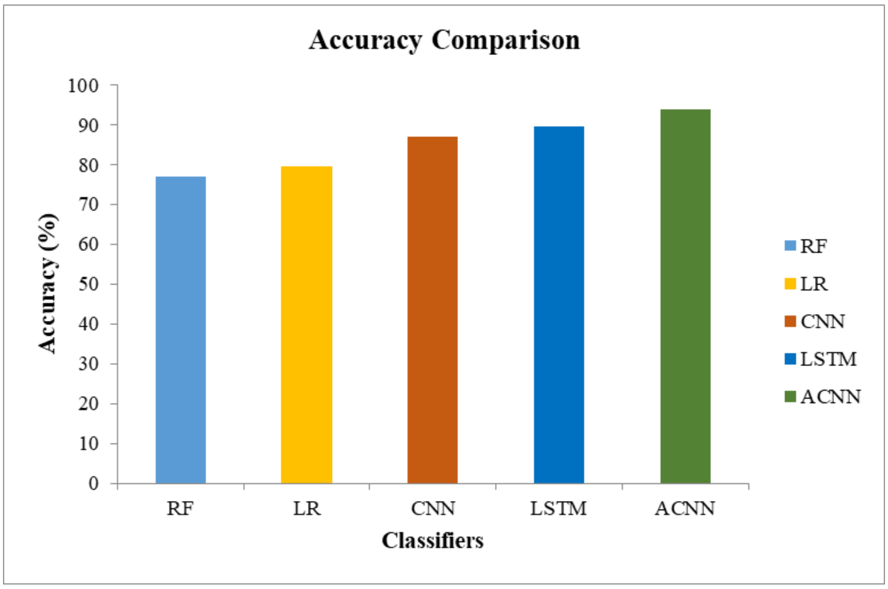

Figure 4: Accuracy Performance Comparison in Various Classification Methods

Figure 4 illustrates the accuracy comparison of the introduced ACNN with existing approaches, including RF, LR, CNN, and LSTM in various classification techniques. Compared to other current approaches, the introduced method may achieve a high accuracy rate, as shown in Figure 4. It is the best method for predicting supermarket sales, with a high accuracy rate of 93.9%. The accuracy rates of the existing methods, such as RF, LR, CNN, and LSTM, are 76.9%, 79.5%, 86.9%, and 89.5%, respectively, which is lower than the accuracy rates of the introduced method.

5. Conclusion and Future Work

This study reviews and thoroughly discusses predictions of supermarket sales in retail stores using machine learning (ML) and deep learning (DL) techniques. Preprocessing, feature selection, and classification are the three primary processes in ML and DL-based approaches primarily studied for SSP in retail stores. Preprocessing, feature selection and classification, merits and demerits, some metrics, datasets, and classification results are analysed for SSP in every single piece of extant literature. The review paper also provided ML- and DL-based techniques for resolving SSP issues in retail establishments. It also fully discusses the simulation outcomes of several DL-based approaches in retail establishments. Data with complex and extensive structures can be learned using DL. In the current assessment, DL was used to build a model that forecasts the growth and decline of sales at a retail location, and its usability was evaluated. The model created using the current research can be used to predict future sales depending on the sales of a specific product on a given day at a specific location of medium size to record the highest sales prediction at a specific location. Eventually, taking into account the prerequisites for applying deep learning to retail enterprises, this research demonstrated the initial stages of an example of the technique for sales prediction. The effectiveness of the proposed approach is assessed using data gathered from the classification-related Rodolfo Saldanha marketing campaign. When picking the supermarket sales prediction in retail outlets, DL techniques consider accuracy the most crucial parameter. In the future, it will have to be used for increasingly larger-scale data and consider broadening the forecasts' applicability. Setting the parameters required for certain sales estimates and introducing optimal values will be crucial to accomplishing this task.

Acknowledgement: The support of all my co-authors is highly appreciated.

Data Availability Statement: This study uses benchmark data available online to conduct the research. This is a fresh study done by the authors.

Funding Statement: There has been no funding obtained to help prepare this manuscript and research work.

Conflicts of Interest Statement: No conflicts of interest have been declared by the author(s). This is the authors' fresh work. Citations and references are mentioned as per the used information.

Ethics and Consent Statement: The consent has been obtained from the colleges during data collection and has received ethical approval and participant consent.

References

1. S. Islam and S. H. Amin, "Prediction of probable back order scenarios in the supply chain using distributed random forest and gradient boosting machine learning techniques," Journal of Big Data, vol. 7, no. 1, pp. 1–22, 2020.
2. G. Verstraete, E.-H. Aghezzaf, and B. Desmet, "A data-driven framework for predicting weather impact on high-volume low-margin retail products," J. Retail. Consum. Serv., vol. 48, no.5, pp. 169–177, 2019.
3. W. R. Padilla, J. García, and J. M. Molina, "Improving time series forecasting using information fusion in local agricultural markets," Neurocomputing, vol. 452, no.9, pp. 355–373, 2021.
4. G. Papacharalampous, H. Tyralis, and D. Koutsoyiannis, "Predictability of monthly temperature and precipitation using automatic time series forecasting methods," Acta Geophys., vol. 66, no. 4, pp. 807–831, 2018.
5. A. Miklosik, M. Kuchta, N. Evans, and S. Zak, "Towards the adoption of machine learning-based analytical tools in digital marketing," IEEE Access, vol. 7, pp. 85705–85718, 2019.
6. S. Islam, S. H. Amin, and L. J. Wardley, "Machine learning and optimization models for supplier selection and order allocation planning," Int. J. Prod. Econ., vol. 242, p. 108315, 2021.
7. R. Odegua, "Applied Machine Learning for Supermarket Sales Prediction". Predictive Machine Learning in Industry, pp. 1–8, 2020.
8. M. Tom, N. Raju, A. Issac, J. James, and R. R. Saritha, "Supermarket Sales Prediction Using Regression," International Journal of Advanced Trends in Computer Science and Engineering, vol. 10, no. 2, pp. 1153–1157, 2021.
9. M. Singh, B. Ghutla, R. L. Jnr, A. F. Mohammed, and M. A. Rashid, "Walmart's Sales Data Analysis-A Big Data Analytics Perspective," in 4th Asia-Pacific World Congress on Computer Science and Engineering, Fiji, 2017, pp. 114–119.
10. H. Wang, "Sustainable development and management in consumer electronics using soft computation," vol. 2019, no. 1, pp. 49–56, 2019.
11. S. Cheriyan, S. Ibrahim, S. Mohanan, and S. Treesa, "Intelligent sales prediction using machine learning techniques," in 2018 International Conference on Computing, Electronics & Communications Engineering (iCCECE), UK, pp. 53-58, 2018.
12. C. Ching-Chin, A. I. Ka Ieng, W. Ling-Ling, and K. Ling-Chieh, "Designing a decision-support system for new product sales forecasting," Expert Syst. Appl., vol. 37, no. 2, pp. 1654–1665, 2010.
13. T.-M. Choi, C.-L. Hui, N. Liu, S.-F. Ng, and Y. Yu, "Fast fashion sales forecasting with limited data and time," Decis. Support Syst., vol. 59, pp. 84–92, 2014.
14. Y. Zuo, K. Yada, and A. B. M. S. Ali, "Prediction of consumer purchasing in a grocery store using machine learning techniques," in 2016 3rd Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE), Fiji 2016, pp. 18-25.
15. J. Yang and S. Ramingwong, "Analysis of Sales Influencing Factors and Prediction of Sales in Supermarket based on Machine Learning Technique," Data Science and Engineering (DSE) Record, vol. 3, no. 1, pp. 66–77.
16. J. Wang, "A hybrid machine learning model for sales prediction," in 2020 International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI), China, 2020, pp. 363-366.
17. R. F. Ali, A. Muneer, A. Almaghthawi, A. Alghamdi, S. M. Fati, and E. A. Abdullah Ghaleb, "BMSP-ML: big mart sales prediction using different machine learning techniques," IAES Int. J. Artif. Intell. (IJ-AI), vol. 12, no. 2, p. 874, 2023.
18. F. Jiménez, G. Sánchez, J. M. García, G. Sciavicco, and L. Miralles, "Multi-objective evolutionary feature selection for online sales forecasting," Neurocomputing, vol. 234, pp. 75–92, 2017.
19. S. F. Pane, A. G. Putrada, N. Alamsyah, and M. N. Fauzan, "A PSO-GBR solution for association rule optimization on supermarket sales," in 2022 Seventh International Conference on Informatics and Computing (ICIC), Indonesia, 2022, pp. 1-6.
20. H. Pan and H. Zhou, "Study on convolutional neural network and its application in data mining and sales forecasting for E-commerce," Electron. Commer. Res., vol. 20, no. 2, pp. 297–320, 2020.
21. Y. Kaneko and K. Yada, "A deep learning approach for the prediction of retail store sales," in 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW), Barcelona, Spain, 2016, pp. 531-537.
22. M. J. Samonte, E. Britanico, K. E. M. Antonio, D. Vega, J. E. J. Espejo, and T. J. P. Samonte, "Applying Deep Learning for the Prediction of Retail Store Sales," in Proceedings of the International Conference on Industrial Engineering and Operations Management, Nsukka, 2022, pp. 185–195.

23. K. Punam, R. Pamula, and P. K. Jain, “A two-level statistical model for big mart sales prediction,” in 2018 International Conference on Computing, Power and Communication Technologies (GUCON), Greater Noida, India, pp.617-620, 2018.
24. L. Alexander and T. Christopher, “An Ensemble Based Predictive Modeling in Forecasting Sales of Big Mart,” International Journal of Scientific Research (IJSR), vol. 5, no. 5, pp. 724–727, 2016.
25. H. Zhaoweijie, “Forecast Rossmann Store Sales Base on XGboost Model, 2nd International Conference on Economic Management and Model Engineering (ICEMME),” Chongqing, China, 2020, pp. 521-525.
26. Y. Niu, “Walmart Sales Forecasting using XGboost algorithm and Feature engineering. International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE),” Bangkok, Thailand, 2020, pp. 458-461.
27. B. K. Jha and S. Pande, “Time series forecasting model for supermarket sales using FB-prophet,” in 5th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2021, pp. 547-554.
28. P. Kaunchi, T. Jadhav, Y. Dandawate, and P. Marathe, “Future sales prediction for Indian products using convolutional neural network-long short term memory,” in 2021 2nd Global Conference for Advancement in Technology (GCAT), Bangalore, India, 2021, pp. 1-5.
29. Y. Dai and J. Huang, “A sales prediction method based on LSTM with hyper-parameter search,” J. Phys. Conf. Ser., vol. 1756, no. 1, p. 012015, 2021.
30. K. Chen, “An online retail prediction model based on AGA-LSTM neural network,” in 2020 2nd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI), Taiyuan, China, 2020, pp. 145-149.
31. Y. Lingxian, K. Jiaqing, and W. Shihuai, “Online retail sales prediction with integrated framework of K-mean and neural network. Pervasive Health: Pervasive Computing Technologies for Healthcare,” Beijing, China, pp. 115–118, 2019.
32. J. Ding, Z. Chen, L. Xiaolong, and B. Lai, “Sales Forecasting Based on CatBoost,” in 2020 2nd International Conference on Information Technology and Computer Application (ITCA), Guangzhou, China, 2020, pp. 636-639.
33. J. Chen, W. Koju, S. Xu, and Z. Liu, “Sales forecasting using deep neural network and SHAP techniques,” in 2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE), Nanchang, China, 2021, pp. 135-138.
34. Kaggle.com. [Online]. Available: https://www.kaggle.com/datasets/rodsaldanha/marketing-campaign. [Accessed: 27-Sep.-2022].