Journal of Artificial Intelligence and Capsule Networks (ISSN: 2582-2012)
www.irojournals.com/aicn/

Enhanced Sales Forecasting in Corporate Cafeterias: An LSTM Approach

Ruban Gladwin M.[1], Jenifer P.[2]

[1]Associate Professor, Department of Electronics and Communication Engineering, Jayaraj Annapackiam CSI College of Engineering, Nazareth, India.

[2]Assistant Professor, Computer Science and Engineering, Francis Xavier Engineering College, Tirunelveli, India.

Email: [1]rubangladwin@gmail.com, [2]jeniferjebavaram@gmail.com

Abstract

A key metric for improving business cafeteria operations, inventory control, and customer satisfaction is accurate sales forecasting. k-Nearest Neighbors (k-NN) and other standard machine learning algorithms usually fail to identify temporal patterns in sales data, which leads to predictions that are not accurate. We provide a better forecasting technique that overcomes this limitation by utilizing Long Short-Term Memory (LSTM) networks, a specific kind of recurrent neural network that can recognize long-term dependencies in sequential data. The performances of LSTM and k-NN were compared and evaluated using a Kaggle corporate cafeteria sales dataset. Based on experimental results, the LSTM model outperformed k-NN with a mean prediction accuracy of 87.57% compared to 83.99%. The significance of the improvement ($p = 0.038$) was validated by statistical testing with an independent samples t-test.

Keywords: Long Short-Term Memory (LSTM), k-Nearest Neighbors (k-NN), Machine Learning, Sales Forecasting, Enhancement, Cafeteria.

1. Introduction

Small cafeterias tend not to have point-of-sale (POS) systems or software that would allow them to record comprehensive sales information and produce useful reports. Their lack

DOI: https://doi.org/10.36548/jaicn.2025.3.002
Received: 10.06.2025, received in revised form: 15.07.2025, accepted: 07.08.2025, published: 22.08.2025

of such systems further fuels inventory management issues since faulty sales predictions result in overstocking, wastage, or shortages, affecting profitability and customer satisfaction. Also, without sound sales projections, budgeting becomes complicated, frustrating wise decisions regarding staffing, buying, pricing, and business strategy.

The incentive for reliable sales forecasting in this area is evolved from the time-sensitive environment of restaurant businesses [1]. Studies on different restaurant environments emphasize that data preparation has a large impact on the applied forecasting method. Although several statistical, machine learning, and deep learning algorithms have demonstrated potential, each of them has certain limitations, as predicted by the 'No Free Lunch' theorem [2].

In today's computational environment, machine learning techniques and neural networks have proven to be strong competitors to conventional statistical analysis for time series prediction. Forecasting sales, which is a time series by nature, entails the prediction of future values using past data [2], [3]. In recent times, Recurrent Neural Networks (RNNs), and more specifically, Long Short-Term Memory (LSTM) networks, have been used extensively to forecast sales with encouraging results [2], [3], [4].

While there are studies that only evaluate classic models like linear regression, decision trees, or support vector machines for restaurant demand forecasting [5], [6], [7], they do not typically involve recurrent neural networks (RNNs). On the other hand, some studies, for instance [8], involve RNN models but without comparison with other statistical or machine learning families. The special power of LSTMs, illustrated by [9], to perform meta-learning upon gradient descent training enables them to update predictions for new data based on inferred hidden states from previous training instances. Such a natural ability makes LSTMs especially suitable for recording rich temporal relationships in sales data.

2. Materials and Methods

Fig.1 depicts the architecture of an LSTM-based corporate cafeteria sales forecasting system. Emphasizing the central role played by Historical Sales Data as input, it begins with the typical Data Preparation step. Such data is painstakingly preprocessed using aggressive data preprocessing techniques like handling missing values, numerical feature scaling, and categorical data encoding to provide coherency and purity. After processing the data, feature engineering builds informative data by combining weather and other environmental factors that are most likely to exert a strong influence on cafeteria sales, special date detection (festivals, corporate events), and building time features such as the day of the week. Clean data is again divided into Training, Validation, and Test Sets to facilitate building appropriate models and estimation.

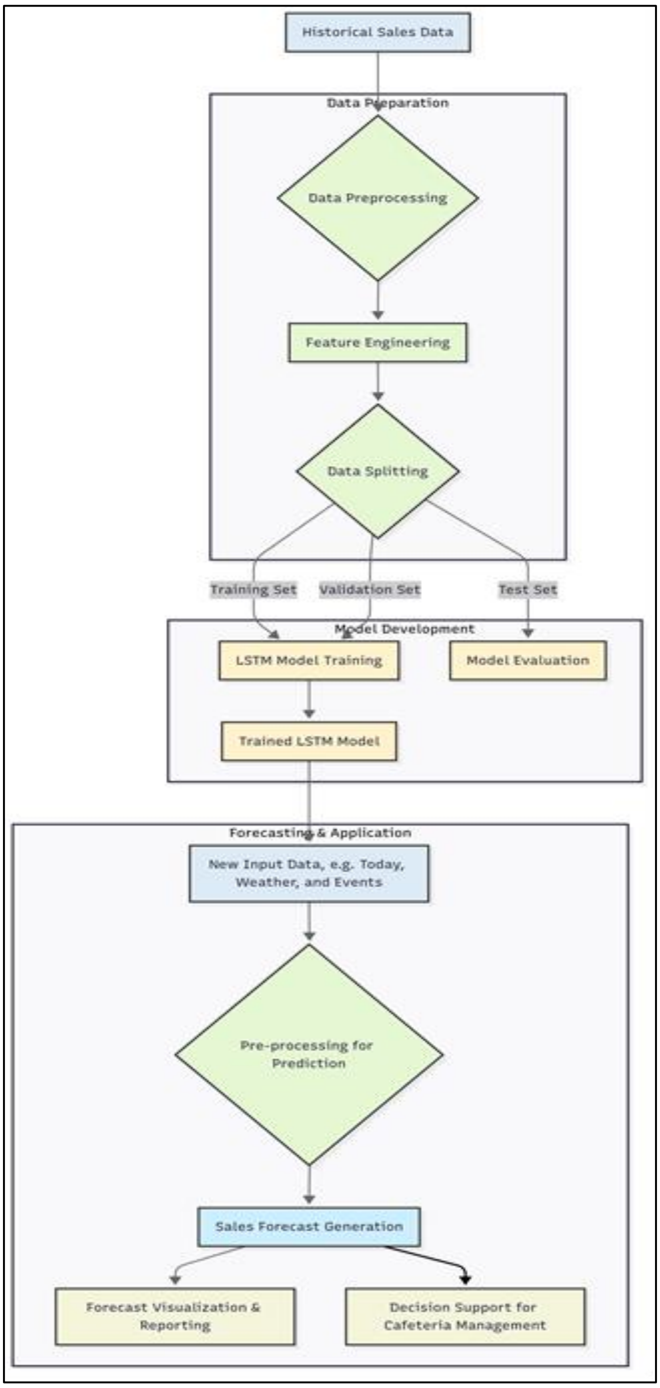

Figure 1. Flow Chart of LSTM-Powered Sales Forecasting for Corporate Cafeterias

The shared goal of the second primary phase, model construction, is to build and evaluate the system's prediction ability. The resulting patterns and lasting relationships in the sequential history of sales are determined by the Training Set as the input to the LSTM Model Training. In this learning process, the Validation Set is used to the maximum extent to adapt hyperparameters and avoid overfitting so that the model is at its best when it experiences new data. Then, the model's performance is precisely measured in Model Evaluation on an independent Test Set after it has been trained. Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) are just a few of the metrics used to gauge the accuracy of the model. An optimized and certified Trained LSTM Model is the output of this phase, which can now be implemented in real life.

The final stage, Forecasting & Application, shows how the model's training already has practical uses in cafeteria management. The system receives fresh input data, such as current and upcoming pertinent facts like dates, forecasted weather, and upcoming business events. It's noteworthy that, due to similarities, the new data is pre-processed in the same manner as the training input. Following pre-processing, the input is fed into the Sales Forecast Generation process, where the LSTM model generates precise forecasts of future sales. Forecast Visualization & Reporting then provides stakeholders with timely information based on those forecasts In order to maximize overall operating efficiency and profitability, the system also offers cafeteria management informative decision support. This help enables maximum inventory, efficient staff management, efficient menu planning, and targeted promotion.

A. Datasets:

The Corporate Restaurants Sales dataset used in this research was sourced from Kaggle [18] and stored in the .csv file format. The information contains 1111 instances and 37 features (as per the abstract), which are diverse attributes used for sales prediction. In order to conduct experiments, the information was divided into test and train. That is, 80% was kept for training and 20% for testing for the Novel LSTM and k-NN models. To enable strong testing and statistical contrast, both models were performance-tested on 10 independent runs. Data were randomly split into the 80/20 train-test splits per run. This enabled the computation of 10 estimates of accuracy per algorithm, with two sample sets (N=10 for LSTM, N=10 for k-NN) to be employed in further statistical analysis. All tests were run on Python 3.6

B. Long Short-Term Memory:

The LSTM network employs a gated cell to store data, akin to computer memory. No other neural networks learn to read and write data from earlier time steps except the cells of the LSTM network [11], [15]. One of the activities of an LSTM network resembles computer memory because it utilizes cells to store and maintain information [17].

Formula:

$Ct = ft \times Ct - 1 + it \times Ct'$

Where:

- Ct is the current memory cell value.
- t is the forget gate value.
- Ct−1 is the previous memory cell value at time step t−1.
- it is the input gate value.
- Ct′ is the candidate cell value.

Algorithm for LSTM:

Step 1: Import Necessary Libraries.

Step 2: Preprocess Data into Sequences and Targets.

Step 3: Split Data into H_train, H_test, R_train, R_test.

Step 4: Utilize train_test_split() for Data Splitting.

Step 5: Construct LSTM models with Layers.

Step 6: Compile Model.

Step 7: Train LSTM Model using Training Data.

Step 8: Evaluate Model using Test Data.

Step 9: Make Predictions and find accuracy

C. k-Nearest Neighbors:

One supervised learning regression and classification algorithm is referred to as k-Nearest Neighbors (k-NN). Decision in k-NN classification are achieved by the majority of the object's k nearest neighbors in the feature space. The output in regression tasks is the weighted sum or average of the k nearest neighbors. Ultimate categorization relies on the proximity of the test set to the training set [14].

Formula:

Predicted Class=argmaxc∑i=1kI(yi=c)

- yi is the class label of the i-th nearest neighbor.
- I(yi=c) is an indicator function that equals 1 if yi=c (where c is the class being considered) and 0 otherwise.

Algorithm for k-NN:

Step 1: Import necessary libraries.

Step 2: Convert the extracted features from the datasets into numerical values.

Step 3: Split the dataset into training and testing sets using train-test split.

Step 4: Choose the value of k (the number of nearest neighbors to consider).

Step 5: Create a k-NN classifier object using the imported k-NN algorithm.

Step 6: Train the k-NN classifier using the training data.

Step 7: Use the trained classifier to make predictions on the test data.

Step 8: Evaluate the model's performance.

Step 9: Compute accuracy

3. Statistical Analysis

Using SPSS software, statistical comparisons of the performance of the k-NN and LSTM algorithms were computed. Group statistical measures are shown in Tables 3 and 4 and reflect variations in the calculated algorithms as statistically different. The data were partitioned into 80% training and 20% testing datasets of 10 runs each, as given in the

"Datasets" subsection [10]. To find the statistical difference between the accuracy of the two algorithms, an independent samples t-test was used. Here, the mean accuracy values of the LSTM and k-NN tests are being compared. SPSS also computed the standard error of the mean and the standard deviation. Note that the "Accuracy" attained in each iteration round is the dependent variable for statistical analysis, and the "Algorithm Type" (k-NN or LSTM) is the independent variable.

4. Results

This section presents a statistical comparison of the two algorithms, K-Nearest Neighbors (k-NN) and Long Short-Term Memory (LSTM). In contrast to k-NN which attained a mean accuracy of 83.99%, the LSTM continuously showed a higher mean accuracy of (87.57%), as shown in Table 3. Additionally, there was less variation in LSTM's performance throughout the runs, as evidenced by the fact that its standard error of the mean (1.05411) was lower than that of k-Nearest Neighbors' (1.20216).

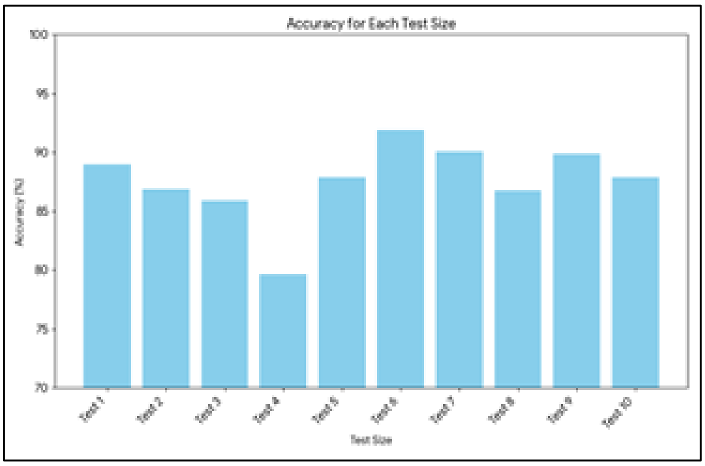

Figure 2. Accuracy for Each Test Size – LSTM Approach

The accuracy percentages for ten distinct LSTM approach tests are shown in graph Fig. 2. From a low of 79.61% (Test 4) to a high of 91.93% (Test 6), the accuracy values vary. The majority of tests have accuracies in the low 90s and mid-to-high 80s. With Test 6 being the most accurate and Test 4 being the least accurate, this visualization makes it easy to compare performance across the different tests. Table 1 shows the precise accuracy scores for each of the ten LSTM evaluation runs, which resulted in an average accuracy of 87.57%. The accuracy scores for each of the ten k-Nearest Neighbors evaluation runs are also shown in Table 2, which yield an average accuracy of 83.99%.

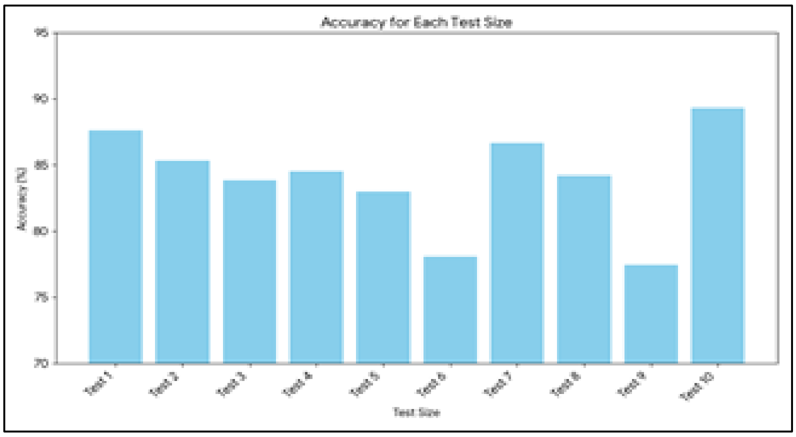

Figure 3. Accuracy for Each Test Size – K-NN Approach

Figure 3 shows the accuracy performance of the K-NN approach across ten distinct tests. The accuracy values fall between roughly 77% and 90%. Test 9 has the lowest accuracy (77.42%), while Test 10 has the highest (89.31%). In general, the accuracies fall within the mid-80s, suggesting that performance varies depending on the test scenario. The statistical significance of the variation in mean accuracies was evaluated using an independent samples t-test, as indicated in Table 4. A p-value of 0.038 ($p < 0.05$) was found by the test. This finding shows that LSTM performs noticeably better than k-Nearest Neighbors, and that the observed accuracy difference between the two algorithms is statistically significant. The mean accuracies and standard deviation errors of the two algorithms are graphically compared in Figure 1.

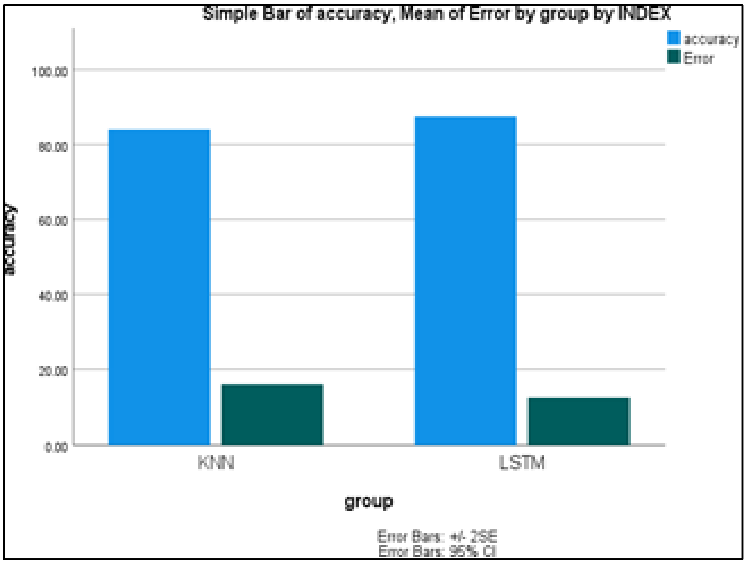

Figure 4. Comparison of LSTM and KNN in Terms of Accuracy

Fig. 4 shows the comparison of Long Short Term Neural Networks and k-Nearest Neighbors in terms of accuracy. The mean accuracy of Long Short Term Neural Networks is better than that of k-Nearest Neighbors, and the standard deviation is slightly lower than that of k-Nearest Neighbors. X-axis: LSTM vs kNN. Y-axis: Mean accuracy of detection ±2 SD.

Table 1. Accuracy of Corporate Cafeteria Sales Prediction using Long Short-Term Memory for 10 samples out of 30 (Accuracy = 87.57%)

Test Size	**Accuracy**
Test 1	88.96
Test 2	86.83
Test 3	85.92
Test 4	76.61
Test 5	87.91
Test 6	91.93
Test 7	90.07
Test 8	86.73
Test 9	89.86
Test 10	87.91

Table 2. Accuracy of Corporate Cafeteria Sales Prediction using k-Nearest Neighbors for 10 samples out of 30 (Accuracy = 83.99%)

Test Size	**Accuracy**
Test 1	87.63
Test 2	85.34
Test 3	83.86

Test 4	84.52
Test 5	82.96
Test 6	78.07
Test 7	86.65
Test 8	84.21
Test 9	77.42
Test 10	89.31

Table 3. Group Statistic Analysis

Algorithm	N	Mean	Std. Deviation	Std. Error Mean
Accuracy LSTM	10	87.5730	3.33340	1.05411
Accuracy K-NN	10	83.9970	3.80158	1.20216

In Table 1, it is shown that the accuracy of Corporate Cafeteria sales prediction using Long Short-Term Memory for 10 samples out of 30 is 87%. In Table 2, it is shown that the accuracy of Corporate Cafeteria sales prediction using k-Nearest Neighbors for 10 samples out of 30 is 83%. In Table 3, the group statistical analysis represents Long Short-Term Memory (mean accuracy 87.57%, standard deviation 3.33340) and k-Nearest Neighbors (mean accuracy 83.99%, standard deviation 3.80158). Table 4 shows the independent sample t-test: LSTM is statistically significantly better than k-Nearest Neighbors with a p-value of 0.038 ($p < 0.05$). A comparison of Long Short-Term Memory and k-Nearest Neighbors in terms of accuracy is shown in Fig. 2.

Table 4. Independent Sample T-test

Accuracy	Levene's Test for equality of variances		T-Test for Equality of Means					
	F	**Sig.**	**t**	**df**	**sig.(2-tailed)**	**Mean Diference**	**Std. error difference**	**95% conf.Interval Lower**
Accuracy Equal variance assumed	0.21	0.652	2.237	18	0.038	3.57	1.59	0.216
Accuracy Equal variance not assumed			2.237	17.698	0.038	3.57	1.59	0.212

5. Discussion

These observations justify the way in which deep learning models, particularly LSTMs, treat time-series data. Traditional machine learning models like k-NN overfit data without temporal interaction warming-up, as seen in earlier work [2], [3]. However, as a proposed innovation, LSTMs are much better able to extract sequential and long-range interactions from data to account for heterogeneity in sales data, such as seasonality, promotions, and day-of-the-week. LSTMs are particularly well-suited for time-series forecasting in that they can remember and store data from earlier time steps, something that cannot be accomplished by distance-based methods like k-NN, which simply get flattened in either direction. [12] also suggested adding an LSTM model as a technique for sales forecasting, with the issue being hyper-parameter tuning to obtain maximum possible accuracy.

Our findings reaffirm the discovery that good models, on average, perform well for intention demand prediction in dynamic settings like corporate cafeterias. This research adds to the body of literature in the field by providing empirical confirmation of the utility of adopting LSTM over k-NN in optimizing prediction performance in this particular niche of sparsely populated historical data and the complex external inputs typical of cafeteria installations overall. Experimentation would favor LSTM over the k-NN algorithm.

6. Conclusion

This research shows that LSTM networks perform better than traditional machine learning methods like k-NN in forecasting firm cafeteria revenues. The LSTM model performs statistically better ($p = 0.038$) and has an accuracy rate of 87.57% in finding temporal patterns and intricacies in the trend of sales in cafeterias. The results depict the advantage of using deep learning for sales forecasting, including measurable improvements in budget planning, reduced stock loss, and improved resource distribution. Although the current study confirms that LSTM performs on the input data set, a number of opportunities to be investigated in future research on hybrid models, parameter optimization, and the addition of additional context variables such as promotions, holidays, or real-time external variables. Lastly, the suggested approach shows how deep learning may significantly impact the foodservice operation decision-making process, and thus constitute a new practice and field.

References

[1] Schmidt, Austin, Md Wasi Ul Kabir, and Md Tamjidul Hoque. "Machine learning based restaurant sales forecasting." Machine Learning and Knowledge Extraction 4, no. 1 (2022): 105-130.

[2] Lasek, Agnieszka, Nick Cercone, and Jim Saunders. "Restaurant sales and customer demand forecasting: Literature survey and categorization of methods." International Summit, Smart City 360° (2015): 479-491.

[3] Bandara, Kasun, Peibei Shi, Christoph Bergmeir, Hansika Hewamalage, Quoc Tran, and Brian Seaman. "Sales demand forecast in e-commerce using a long short-term memory neural network methodology." In International conference on neural information processing, Cham: Springer International Publishing, (2019): 462-474.

[4] Carbonneau, Real, Kevin Laframboise, and Rustam Vahidov. "Application of machine learning techniques for supply chain demand forecasting." European journal of operational research 184, no. 3 (2008): 1140-1154.

[5] Tanizaki, Takashi, Tomohiro Hoshino, Takeshi Shimmura, and Takeshi Takenaka. "Demand forecasting in restaurants using machine learning and statistical analysis." Procedia CIRP 79 (2019): 679-683

[6] Sanjana Rao, G. P., K. Aditya Shastry, S. R. Sathyashree, and Shivani Sahu. "Machine learning based restaurant revenue prediction." In Evolutionary Computing and Mobile Sustainable Networks: Proceedings of ICECMSN 2020, Singapore: Springer Singapore, (2020): 363-371.

[7] Sakib, SM Nazmuz. "Restaurant sales prediction using machine learning." In Handbook of research on AI and machine learning applications in customer support and analytics. IGI Global, (2023): 202-226.

[8] Stergiou, Konstantinos, and Theodoros E. Karakasidis. "Application of deep learning and chaos theory for load forecasting in Greece." Neural Computing and Applications 33, no. 23 (2021): 16713-16731.

[9] Hochreiter, Sepp, A. Steven Younger, and Peter R. Conwell. "Learning to learn using gradient descent." In International conference on artificial neural networks, Berlin, Heidelberg: Springer Berlin Heidelberg, (2001): 87-94.

[10] Rácz, Anita, Dávid Bajusz, and Károly Héberger. "Effect of dataset size and train/test split ratios in QSAR/QSPR multiclass classification." Molecules 26, no. 4 (2021): 1111.

[11] Cortez, Bitzel, Berny Carrera, Young-Jin Kim, and Jae-Yoon Jung. "An architecture for emergency event prediction using LSTM recurrent neural networks." Expert Systems with Applications 97 (2018): 315-324.

[12] Dai, Yun, and Jinghao Huang. "A sales prediction method based on lstm with hyper-parameter search." In Journal of Physics: Conference Series, vol. 1756, no. 1, p. 012015. IOP Publishing, 2021.

[13] Chniti, Ghassen, Houda Bakir, and Hédi Zaher. "E-commerce time series forecasting using LSTM neural network and support vector regression." In Proceedings of the international conference on big data and Internet of Thing, (2017): 80-84.

[14] Hu, Li-Yu, Min-Wei Huang, Shih-Wen Ke, and Chih-Fong Tsai. "The distance function effect on k-nearest neighbor classification for medical datasets." SpringerPlus 5, no. 1 (2016): 1304.

[15] Kumar, Jitendra, Rimsha Goomer, and Ashutosh Kumar Singh. "Long short term memory recurrent neural network (LSTM-RNN) based workload forecasting model for cloud datacenters." Procedia computer science 125 (2018): 676-682.

[16] Rácz, Anita, Dávid Bajusz, and Károly Héberger. "Effect of dataset size and train/test split ratios in QSAR/QSPR multiclass classification." Molecules 26, no. 4 (2021): 1111.

[17] Samonte, P. M. J., Elixeline Britanico, Karmin Eda Mae Antonio, J. E. J. De la Vega, Tia Julienne P. Espejo, and Danielle C. Samonte. "Applying deep learning for the prediction of retail store sales." In Proceedings of the International Conference on Industrial Engineering and Operations Management, vol. 10. 2022.

[18] https://www.kaggle.com/c/restaurant-revenue-prediction/data.