Special Event Item Prediction System for Retails – Using Neural Network Approach.

Tilshi Alwis

Department of Computer Science, Faculty of Computing, General Sir John Kotelawala Defence University , Ratmalana, Sri Lanka.
35-cs-0018@kdu.ac.lk

Punsisi Pemarathna

Department of Computer Science, Faculty of Computing, General Sir John Kotelawala Defence University , Ratmalana, Sri Lanka.
punsisi@kdu.ac.lk

ABSTRACT

Selling and buying is the general process marketing field follows. Nowadays marketing field bonded with the modern technology, and it highly effected to field expandability. Marketing become fruitful when it achieves its key points which are called sales and profit. Mostly people are move to the retails because all the essentials and other things can buy from one place. There are many technological concepts involve with marketing field as an enhancement. Prediction processes, data analysis, item designing and profit calculation are some representatives for those concepts. This study is a prediction process, developed for retails using machine learning approaches. Item sales data analyzed and generated prediction results on set of items which are given maximum or expected profit margins and which items satisfy the customer most. Item suppliers are key stakeholder type a retail can have, there is a recommender system in this approach for suppliers and the recommendation is based on past sales data. There are certain types of machine learning approaches used in sales item prediction, sales item feature prediction, sales price prediction and etc. Novelty of this research is, it focused only special event items such as items in Christmas season, items specialized for Mother's Day, Valentine Day, Sinhala, and Tamil new year and etc. This research process had completely followed the machine learning neural network concept. Recurrent Neural Network is subpart of neural networks and this research study followed up through this RNN method. Neural network had applied using a form of machine learning called deep learning. This model had worked on sequential data therefor LSTM (Long Short-Term Memory) layers were used and to avoid overfitting issue several dropout layers were used. The results prove neural network method has highest accuracy.

Key words: Sales prediction, Special Event Items, Machine Learning, Neural Network, Deep Learning, Retail.

1   INTRODUCTION

Artificial Neural Networks(ANN) are basically an inspiration of human brain. Mostly it comes under supervised learning approaches in machine learning and Recurrent Neural Networks(RNN) are subdivision of ANN. There for RNN can discuss as a machine learning approach. Machine learning is a specialized technical discipline that may be used to perform a variety of tasks, including data prediction, data mining, data optimization, data categorization, clustering, and dimensionality reduction. It is a rapidly emerging subject of computer science, with applications that are extremely complicated when combined with other technology (Dey, 2016).. Data Prediction is one of them, and machine learning approaches have a big influence on it. Machine learning technology is used in several areas of data prediction, including sales item prediction, item design prediction, sales profit predictions, medical diagnosis prediction using symptoms (Cancer, Virus, and so on), bankrupt's predictions, personality prediction using textual data, item popularity prediction (Ex; Car popularity prediction), and so on. This

method of prediction calls for the use of previously examined data. The main goal of this study is to estimate retail sales for things that are sold during special events, days, or seasons. Today, sales items have become a critical component of the corporate world, and sales item data is rapidly increasing. Clothing retails, vegetable and fruit retails, retails with all needs, cosmetic retails, and so on are all examples of retails. Two categories of retails are investigated in this study: retails that solely sell gift items and retails that sell both essential and gift items. Retailers or retailers earn on sales items.

Predicting sales items is a critical procedure nowadays. It helps to meet customer expectations by increasing sales profit, reducing cost overruns, and increasing customer satisfaction. Customer expectations are constantly being updated with new features. In stores, there is an item management team that is responsible for identifying and analyzing customer expectations. People tend to buy everything in one spot due to their hectic lifestyles. "Arpico Super Center" and gift shops like "Vondy Party City" are two of Sri Lanka's most well-known merchants. Items that fall within the categories of exceptional events, seasonal festivals, and calendric days are considered in this study. For each of the dates listed above, there are a variety of specific gift items available, such as cakes, chocolates, flowers, ornaments, teddy bears with varying colors, shapes, tastes, amount of flowers in the bouquet, band, and so on. Calendar days, seasonal festivals, and special functions are the three types of special events that are relevant to retailers.

Originality of this research study is, it is considered special event sales items. There are three types of special events in the retail industry. Calendric day: These are events that occur on a specific date or month, such as Mother's Day, Father's Day, Children's Day, Lovers' Day, and so on. Most likely, it is an international holiday, although it can also be limited to a single country, as in Sri Lanka's May Day and Independence Day. The majority of individuals observe these event days as a personal preference, tradition, or habit. They used to exchange gifts or special mementos during such gatherings. Retailers will be able to enhance their sales and profit margins during that time period by offering things related to a specific event. Retailers can also gain extra benefits by preparing for upcoming events utilizing item prediction processes based on previous year's data. For such process, the proposed "Special Event Item Prediction System" will be quite valuable. Seasonal festivals: Seasonal festivals can be classified as a different form of calendar event because they occur on a specific date or month throughout the year. Seasonal festivities include "Christmas," "Vesak," "Ester," "Ramazan," and "Sinhala and Tamil New Year," to name a few. During this occasion season, people are extremely motivated to purchase gift items, decorations, and meals. Then these particular event types will provide an opportunity to take advantage of the benefits indicated above(in the calendric day paragraph). Seasonal festivals, in particular, are a form of shopping occasion. Retailers should have prepared their specific stocks by then. Special functions do not have a unique or specified date or month in the calendar year. The most common instances of special functions are birthdays and wedding anniversaries. The "retail event management team" should pay greater attention to such events because they can have deals on certain things every day. As previously stated, by arriving early to these types of events, retailers may meet customer expectations while maintaining predicted profit margins.

In the proposed work, we present a machine learning subdivision called RNN-based special event item prediction system for retailers. RNNs are used to perform item recommendation and prediction tasks. This approach will aid in increasing predicted profit margins and meeting consumer expectations in certain events. By researching existing systems, we were able to determine the system's critical importance and requirements. As a consequence, the appropriate algorithms and machine learning techniques were identified. The suggested system's additional goals are to prepare the dataset, design the systems, implement the system, train the model, and assess the outcomes.

2   RELATED WORKS

Predicting future customer purchases is very important and support to planning the inventory of retail, shop, or warehouse (ndr«es Mart«õnez, Claudia Schmuck, Sergiy PereverzyevJr., Clemens Pirker, Clemens Pirker, Markus Haltmeier, 2018). In paper "A Machine Learning Framework for Customer Purchase Prediction in the Non-Contractual Setting" proposed an advanced analytics tools to perform above mentioned task. Their proposed application implemented through various machine learning algorithms for binary classification. They had used three types of classification methods called: logistic

Lasso regression, extreme learning machine and gradient tree boosting. These methods are totally different one from another, reason to use such methods is to increase accuracy with reasonable computational effort. From the results they had proved gradient tree boosting has highest accuracy. This prediction done for before one month to get the inventory for next month.

Sales time series forecasting has been taken by a system using Stacking approach for machine learning models (Pavlyshenko, 2018). Stacking approach provide ability to use the results of multiple model prediction on the validation set as input regression for next level model and also this approach caused to improve accuracy validation and out of sample data sets. The proposed solution is a three-level model and as the first level used single models, mostly XGBoost machine learning algorithm. As the second stacking level used three models called Extra Tree and Neural Network models. Then summed the results of second level with weights of third level.

There are large number of sales forecasting researchers adopted by neural network models. As the paper (A.L.D. Loureiro, V.L. Miguéis, Lucas F.M. da Silva, 2018) includes, Decision Tree, Random Forest, Support Vector Regression, Artificial Neural Network and Linear Regression are the machine learning techniques they used. For achieve better performance they had used deep learning approach. From evaluation matrices proved Random Forest method has accurate results and better performance than other techniques. This proposed model developed with 10 variables, there for products were characterized by considering those variables. Product price is the only numerical variable this approach contained.

In paper "Application of Long Short-Term Memory Neural Network to Sales Forecasting in Retail" described a case study on sales forecasting in using historical data. The methodology had followed up deep learning based method called Long Short-Term Memory(LSTM). LSTM is a sort of Recurrent Neural Network (RNN) (Yu Q., Wang K., Strandhagen J.O., Wang Y., 2018). In this research study forecasting had done using sales data of 66 products sold in 45 weeks. Most important fact is forecast had done in week level, there for sales dates of first 30 weeks used as training data, the second 15 weeks used as test data. As common way sales values are scaled between [0,1].

Kui Zhao and Can Wang had presented a sales forecasting approach using Convolutional Neural Network(CNN) This study presented for overcome the limitations of existing sales forecasting methods. There for this approach learn effective features automatically form structured raw log data using CNN. In this paper CNN model architecture consists with data frame, convolutional feature maps, activation function, pooling, multiple feature map and fully convention. Finally applied the Linear Regression to obtain the final results of specified feature vector. Then can conclude in neural network activation function used to overcome the linearity issue by achieving non-linearity (Kui Zhao, Can Wang, 2017) and also, this presented research work transferred the knowledge obtained from one problem to another using transfer learning and as the optimization function it used the Stochastic Gradient Descent(SGD) algorithm. This approach was compared with other several approaches called ARIMA, FE+GBRTA and DNN. But CNN model had achieved highest accuracy.

Mostly Sales forecasting approaches follows the neural network methods. It is possible to use small dataset with a neural network for sales prediction approaches (Rosa María Cantón Croda, Damián Emilio Gibaja Romero, Santiago Omar Caballero Morales, 2018). Neural network implementation done respectively training, validation and prediction. Sales forecasting can do considering small time slots, in (Rosa María Cantón Croda, Damián Emilio Gibaja Romero, Santiago Omar Caballero Morales, 2018) had considered the month by month. The models used for this is Simple Moving Average(SMA) and Artificial Neural Network(ANN). SMA can calculate approximation error for each time period. But in ANN cannot replicate entire time series behavior, database size is the main reason for this. There for ANN done prediction for a time period which is larger than SMA consider time period. According to (Rosa María Cantón Croda, Damián Emilio Gibaja Romero, Santiago Omar Caballero Morales, 2018) ANN considered three months. When search through results ANN error is less than SMA. Then ANN has best performance.

Yuto and Katsutoshi were presented a deep learning approach for sales prediction and provide deep learning is effective for analyzing the point of sales data of retail stores (Y. Kaneko and K. Yada,, 2016). Prediction had done for particular day by considering past three years' worth point of sales data. Deep Learning L1 regularization and Logistic Regression models were used build up the prediction model and deep learning model achieve more accuracy than other. Most important thing is used attributes were divided into three categories and highest accuracy achieved by category 1 data. Number

of attributes vary form 62 in category 1 to 3312 in category 3. There for number of attributes caused to prediction accuracy.

Artificial Neural Network(ANN) provides an approximation for all non-linear continuous functions and it is applicable for sales forecasting (M. Garetti, and M. Taisch,, 1999). There are not any unique types of approaches for sales forecasting there for ANN consists with different structures and those structures different from problem to problem. There is a research study for electricity demand forecasting used hybrid structure and data presented in seasonal effects (N. An, W. Zhao, J. Wang, D. Shang, and E. Zhao, 2013). This approach architecture consists with two steps. First step used to reduce the seasonal effects using empirical mode decomposition and second step obtained the demand forecast for specified following time periods considering previous time periods. Multi output feed forward ANN technique used for second step. There is another study of ANN used with stochastic connections for noisy data prevail in water consumptions (H. Rodriguez, V. Puig, J. J. Flores, and R. Lopez,, 2016). Both research studies above mentioned are constructed with ANN architecture used for a long time series.

In paper "House Price Prediction Using Machine Learning and Neural Networks": Author has done extensive study on Predicting housing prices with real factors (Ayush Varma, Abhijit Sarma, Sagar Doshi, Rohini Nair, 2020).The results of research proved that this approach provides minimum error and maximum accuracy than individual algorithms applied It considered parameters are 'square feet area', 'no. of Bedrooms', 'No of Bathrooms', 'Type of Flooring', 'Lift availability' ,'Parking availability' and 'Furnishing condition'. As a unique approach to increase accuracy, used that the actual real estate value also depends on nearby local amenities such as railway station, supermarket, school, hospital, temple, parks etc. Author has done this study using number of algorithms such as Linear Regression, Forest Regression, Bootstrap Regression and Neural Network.

3   METHODOLOGY

As initial data set, past sales data were used. CSV file consisted with several rows and for prediction process used few of them. This proposed application has two types of users called. Special event item suppliers and system admin. There for system functional requirements can divide considering system users and prediction processes,

1. Retail Special Event Item Management Team or Manager (Admin)
    *   Predict the items with higher number of sales (User expected items).
    *   Predict the items which give maximum profit in specific event.
2. Item Suppliers
*   Recommend items or item categories, which can give maximum profit to suppliers by supplying the items to the retail.

There for there are two prediction processes, item prediction for higher number of sales, it based on Sold Quantity attribute of dataset and item prediction for find which items give maximum profit in specific event based on Unit Profit attribute of dataset. Item Recommendation process is same as item prediction for higher number of sales processes.

Methodology of built RNN model has several steps, finalize the data set, finalize the dataset attributes used to prediction processes, RNN model build from Spyder and test the build models accuracies.
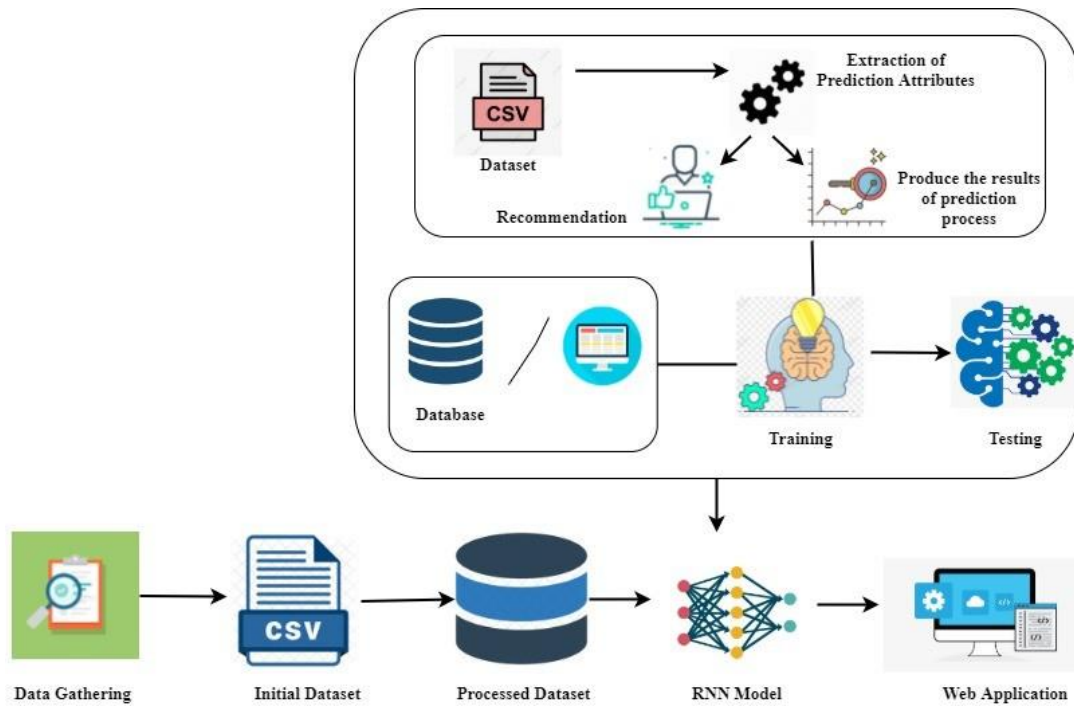
Figure 1 RNN Model Implementation Process.

3.1   Dataset

Table 1 Dataset Attributes

Event Name	Event Id	Item Name	Item Id	Sold Quantity	Unit Price	Unit Profit

Above table columns represented the attributes of referred dataset. Event Id and Item Id used to match the each and every item with the specific event. Sold Quantity, Unit Price and Unit Profit attributes are used for item prediction and item recommendation processes. Dataset is found from Kaggle website and edited with real time sales dataset.

- Unit Price : Used for price forecasting, it would be useful for select the number items which can change the prices in specific time period.

- Unit Profit : Used to predict the profit of items which are selected from price forecasting process.

- Sold Quantity : Used for Predict the items with higher number of sales (User expected items) and Recommend items or item categories, which can give maximum profit to suppliers by supplying the items to the retail.

3.2   Recurrent Neural Networks.

In neural network approaches, RNN is the most prevalent type. RNNs basically take the output from the previous step/layer and feed it into the next step/layer as input. In most neural networks, all of the inputs and outputs are independent of one another, while RNN notions are the polar opposite. If this research study is correct, while predicting products based on prior sales data, the model must remember the past sales data, and to tackle this problem, the model must contain hidden layers. Hidden states are used in the RNN process to remember important information about the sequence (Yu, Y., Si, X., Hu, C. and Zhang, J., , A review of recurrent neural networks: LSTM cells and network architectures). Because

RNN memory can retain all information about completed computations, it prefers parameters for each input to execute the same task on all inputs/hidden layers in order to create the output. As a result, the complexity of parameters was reduced. This is one of the fundamental distinctions between RNNs and other artificial neural networks.

3.3 Model Built from Spyder.

For the sequential model, the Keras deep learning package was employed and this is a python-based deep learning library that is open-source software. Other deep learning libraries can be used as the frontend with this library type. Keras is a high-level API model, thus it's simple to grasp and use (Gulli, A. and Pal, S., , 2017). The advancement of RNN model implementation was heavily influenced by tensor flow and keras. Each layer of the model has its own activation function, as well as the possibility to add layers to the model. Both of Spyder's prediction models go through a training phase. All of the layers in the network were altered after the training process was done. When it comes to the ultimate accuracy rates obtained during the training, validation, and testing phases. LSTM layers were used extensively in both prediction model layers because it is capable with memorize important information in the dataset and feed data back to the neural network. Relu was employed as the activation function for all LSTM layers in both prediction processors. Both prediction models used the Dense layer as the final layer, with the SoftMax activation function as the Dense layer activation function.
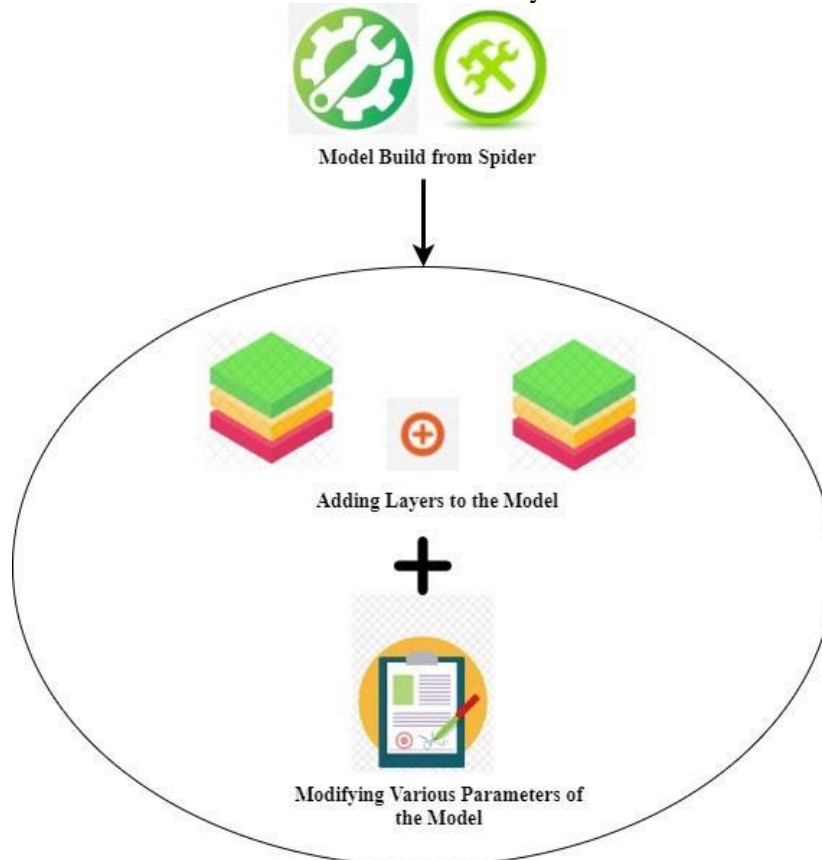

Figure 2 Design the model built from Spyder

Python is the core programming language of this approach and it is the choice for developing prediction and classification models. This is the core programming language for machine learning problems and is classified as a general-purpose computer language. Python is a fairly simple language to learn and use. Python has a number of advanced machine learning libraries (Srinath, K.R., , 2017. ,).

3.4 Changing Layers.

Layers changing process can explain with following scenarios.
1. Considering the inputs shapes changing 50 LSTM layers in Return Sequence.
2. Changing the 50 LSTM layers in Return Sequence consists with 0.2 Dropout and constant 32 batch size.
3. Changing the 50 LSTM layer s without Return Sequence consists with 0.2 Dropout and constant 32 batch size.
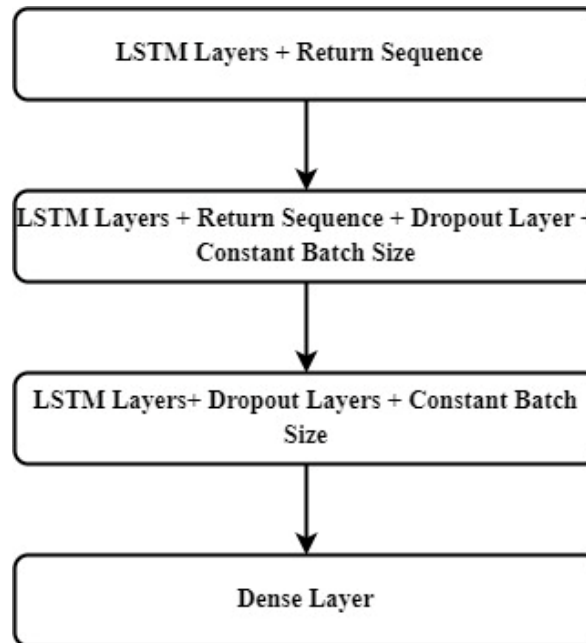4. Finally go through Dense layer and it is the final output layer.

Figure 3 Network architecture of the model developed from Spyder.

3.5 Simulation Setup

Dropout Regulization

The Dropout regulization was set to a constant value of 0.2 over all 50 LSTM layers. It resulted in the creation of an accurate model by preventing the issue of model overfitting.

Number of Epochs

The number of epochs was also a fixed value, there for 25 epochs were used to train the network.

Batch Size Parameter

Batch size also considered with a fixed value; it is 32.

4 RESULTS.

The RNN model was tested under three approaches.

Checking the Accuracy and Loss of RNN model using training dataset.

When the model is being trained, a training accuracy is kept understanding the accuracy acquired by the model during its training phase. This value normally rises with the number of epochs as the model sees more of the training sales dataset.

Checking the Accuracy and Loss of RNN model using validation dataset.

The model examines the sales data in the validation dataset after each epoch to see how accurate the model's weights are for that epoch. After all of the sales data in the validation dataset is input into the model to determine a validation accuracy, the model modifies its weights accordingly, and then repeats the process of having the training dataset stream through all of its layers and the validation dataset do the same at the end of the epoch. This method will be repeated until the training and validation datasets have passed through the model layers a specified number of times (epochs) (Pawar, K., Jalem, R.S. and Tiwari, V., , 2019).

The outcomes of the training technique are used to create two graphs for each model. One of these was the Accuracy graph, which compares the training and validation accuracy plots. The Loss graph, on the other hand, displayed the training loss plot vs the model's validation loss plot.
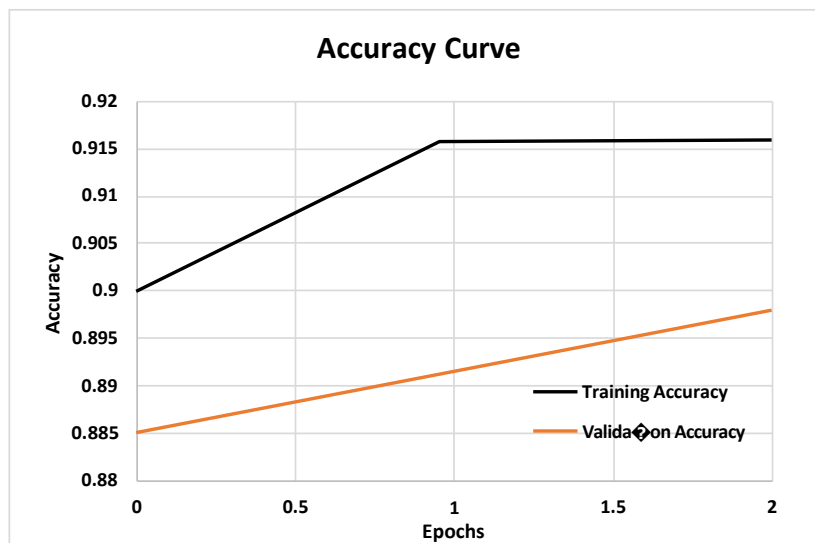

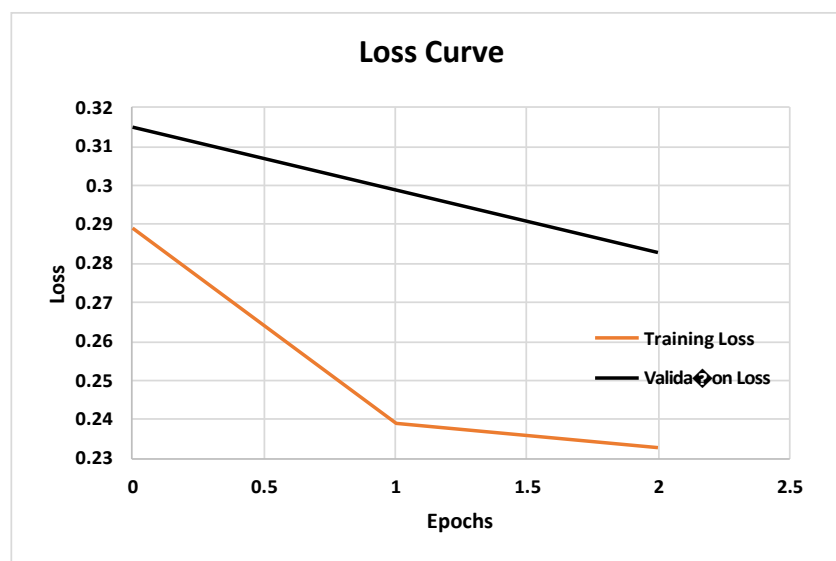Figure 4 Training and Validation Accuracy Graphs for Model Trained using Spyder.

Figure 5 Training and Validation Loss Graph for Model Training using Spyder

s

Checking the Accuracy and Loss of RNN model using testing dataset.

At the end of the training operation, the testing dataset is utilized to determine an unbiased accuracy, which is known as the testing accuracy. A collection of sales data that has never been viewed by the RNN model is employed to achieve unbiased accuracy. The phase of implementation and training is the most important aspect of the entire study process. RNN may be used to create two sorts of prediction models. One prediction is based on the things' profit, while the other is based on the items' sold quantity. Both prediction processes employed item sales data as a data set. These prediction models were created with the Flask Python framework and trained with Spyder.

While the training process is in progress, training accuracy/loss begins to increase/decrease at a faster pace than validation accuracy, and the training process must be terminated.

5 CONCLUSION

During special occasions such as calendar days and seasonal festivals, most retailers are unable to accomplish their profit margins and meet customer expectations due to unfitted present items or décor. This study used a machine learning methodology to propose a retail special event item prediction strategy. This is one of the most crucial learning outcomes in the study process. If the training process cannot be stopped, the model begins to overfit, resulting in erroneous outputs. The testing phase was used to understand and gain some knowledge about the accuracy of the results provided by the RNN models individually, and one flaw was discovered: when a large number of past sales items are present, the event prediction process takes approximately 4 seconds to produce final results.

This proposed approach can enhance with several requirements as future works. Services, users, and events can all be expanded to make this application more useful. This developed system can only be used by retail members and special event item suppliers, and it can be improved as an e-commerce website by giving services to customers. Furthermore, this application can be improved by providing customers with item customization services for a unique special occasion.

This built RNN model generates prediction results based on prior sales data such as sold quantity and profit, as well as item characteristic parameters such as design, color, and size for the prediction process. It will be more feasible than the currently available model.

In the future, our approach could provide location-based retail recommendations to help retailers find the closest suppliers for specific events. This can be accomplished by offering maps to suppliers during the registration process and to administrators on their home page.

6 AUTHOR BIOGRAPHY

Mrs. Punsisi Pemarathne is a Senior Lecturer/Researcher in Computer Science at the Department of Computer Science, General Sir John Kotelawala Defence University. She has obtained her BSc Computer Systems and Networks and her master's in computer and Network Engineering from Sheffield Hallam University in United Kingdom. She has completed her MPhil in Computer Science from University of Sri Jayewardenepura, Sri Lanka. Her research interests include, Swarm Robotics, Artificial Intelligence, Mobile and Wireless Communications, Network Security, Evolutionary Computing, and Internet of Things.

W.H.T.M. Alwis an undergraduate student in the General Sir John Kotelawala Defence University , Sri Lanka and will be graduating in 2022 with a BSc Hons. In Computer Science. Her research interest is in the field of Machine Learning.

7   REFERENCES

A.L.D. Loureiro, V.L. Miguéis, Lucas F.M. da Silva. (2018). Exploring the use of deep neural networks for sales forecasting in fashion retail. *Decision Support Systems, Doi :10.1016/j.dss.2018.08.010.*

Ayush Varma, Abhijit Sarma, Sagar Doshi, Rohini Nair. (2020). House Price Prediction Using Machine Learning. *IEEE.*

Dey, A. (2016). Machine Learning Algorithms: A Review. *International Journal of Computer Science and Information Technologies,, Vol. 7 (3)*, 1174-1179.

Gulli, A. and Pal, S., . (2017). *Deep learning with Keras. .* Packt Publishing Ltd.

H. Rodriguez, V. Puig, J. J. Flores, and R. Lopez,. (2016). Combined holt-winters and GA trained ANN approach for sensor validation and reconstruction :Application to water demand flowmeters. *2016 3rd Conference on Control and Fault-Tolerant Systems (SysTol), on, pp. 202-207,.*

Kui Zhao, Can Wang. (2017). *Sales Forecast in E-commerce using Convolutional Neural Network.* Retrieved 10 2021, from https://arxiv.org/abs/1708.07946

M. Garetti, and M. Taisch,. (1999). Neural networks in production planning and control,. *Production Planning & Control, 10*, 324-339.

N. An, W. Zhao, J. Wang, D. Shang, and E. Zhao. (2013). Using multi-output feedforward neural network with empirical mode decomposition based signal filtering for electricity demand forecasting. *Energy, 49*, 279-288.

ndr«es Mart«õnez, Claudia Schmuck, Sergiy PereverzyevJr., Clemens Pirker, Clemens Pirker, Markus Haltmeier. (2018). A Machine Learning Framework for Customer Purchase Prediction in the Non-Contractual Setting. *European Journal of Operational Research, 239*(3, 2014 Dec).

Pavlyshenko, B. (2018). Machine-Learning Models for Sales Time Series Forecasting. Lviv, Ukraine: IEEE Second International Conference on Data Stream Mining & Processing (DSMP).

Pawar, K., Jalem, R.S. and Tiwari, V., . (2019). n Emerging Trends in Expert Applications and Security. *Springer*, (pp. 493-503).

Rosa María Cantón Croda, Damián Emilio Gibaja Romero, Santiago Omar Caballero Morales. (2018). Sales Prediction through Neural Networks for a Small Dataset. *International Journal of Interactive Multimedia and Artificial Intelligence(IJIMAI) , 5*(Special Issue on Artificial Intelligence Applications), 35-41.

Srinath, K.R., . (2017. ,). Python–the fastest growing programming language. I). *nternational Research Journal of Engineering and Technology (IRJET), 4*(12), pp.354-357.

Y. Kaneko and K. Yada,. (2016). A Deep Learning Approach for the Prediction of Retail Store Sales. 2016 IEEE 16th International Conference on Data Mining Workshops.

Yu Q., Wang K., Strandhagen J.O., Wang Y. (2018). Application of Long Short-Term Memory Neural Network to Sales Forecasting in Retail—A Case Study. . *Advanced Manufacturing and Automation VII., 451*, 11-17.

Yu, Y., Si, X., Hu, C. and Zhang, J., . (A review of recurrent neural networks: LSTM cells and network architectures). A review of recurrent neural networks: LSTM cells and network architectures. *Neural computation, 7*(31), pp.1235-1270.