Demand Forecasting For Wholesale Store Using Deep Learning Methods

MSc Research Project
Data Analytics

Aryan Rajput
Student ID: X20128088

School of Computing
National College of Ireland

Supervisor: Prof. Hicham Rifai

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Aryan Rajput
Student ID:	X20128088
Programme:	Data Analytics
Year:	2021
Module:	MSc Research Project
Supervisor:	Prof. Hicham Rifai
Submission Due Date:	16/12/2021
Project Title:	Demand Forecasting For Wholesale Store Using Deep Learning Methods
Word Count:	6262
Page Count:	20

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Aryan Rajput
Date:	16th December 2021

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Demand Forecasting For Wholesale Store Using Deep Learning Methods

Aryan Rajput
X20128088

Abstract

Demand forecasting has become an interesting topic among e-commerce domain, and it is very pivotal among warehouses. According to the present market trend, it has become increasingly important to deliver things to customers as quickly as possible and on schedule. It is necessary for warehouses to keep track of the availability and demand of things that clients seem to purchase. Forecasting demand is a way of creating a model that seems to predict the demand for a certain item at any point in the future. The demand forecasting tactics can be linked to maintain a smooth distribution of the products that consumers will buy at certain time in the future. Along with to provide better stock management, stock management's another primary aim is to reduce the total cost of stocked inventories. To complete this task, Seasonal Autoregressive Integrated Moving Average(SARIMA) and Recurrent Neural Network(RNN) has been applied. The outcome of both the models has been compared with mean error. SARIMA performed with a Root Mean Square Error(RMSE) value of 235498.28, whereas for RNN the Mean Absolute Error(MAE) value is obtained as 56.87% and correctness of 43.12%.

Keywords— Demand Forecasting, Seasonal Autoregressive Integrated Moving Average, Recurrent Neural Network, deep learning

1 Introduction

Since the beginning of trading system, supply chain management is playing a vital role to control the flow of supplies. As advancement in e-commerce is touching new heights, many businesses are leaning towards adaptive business development. With the help of big data tools and technologies it has become possible to increase the efficiency of the warehouse stores. By considering the options available and the ways of predicting the demand of a particular product at any given time in future has been made possible by using latest technologies and new machine learning methods.

Supply Chain Management (SCM) is the backbone of many global organizations in today's industry. SCM is based on the management and optimization of processes and systems. From the collecting of raw materials through the dispersion of the end product, it is involved in the whole product life cycle. The first and most important stage for large warehouses is to maintain track of product demand in order to ensure that sufficient supply is met without overstocking or running out of stock. Demand forecasting has vastly improved and grown in recent years, as it is a critical and significant part for many businesses. Because warehouse space is scarce and expensive, it is not viable to load them with merchandise. In addition, clogged warehouses make it more difficult to gather and locate orders. And thus, it makes it even more important to manage inventory regularly and store goods as per their current market trend. Numerous attempts have been made to anticipate demand for warehousing items. Many basic and complicated models have been used to develop a long-term method for forecasting demand.

The Figure 1 encompasses the overall view of a warehouse, and how the workers and managers are worried about the important factors for their warehouses. Inventory management depends upon so many factors which needed to be kept in mind like re-order point, average daily sales, turnover and many more. Automating inventory and forecasting the demand could turn such overwhelming tasks a little bit smoother.

Figure 1: Demand Forecasting

Although demand forecast methodologies have already been implemented among multiple sectors, there is always scope of enhancement, when it comes to determining the most appropriate and efficient technique. As previously said, demand forecasting is the industry standard and holds the value among manufacturing companies. The main motivation of this research is to inspect the effectiveness of machine learning algorithms in estimating retail product demand for a warehouse by taking all of these factors into account. The complete method will be carried out with the use of a publicly available e-commerce order demand dataset. The research is going to be carried out using Seasonal Autoregressive Integrated Moving Average (SARIMA) and Recurrent Neural Network(RNN). A relation is going to be established after a comparison of the outcomes received after implementing both the models.

1.1 Research Question and Objective

The research performed is going to answer the following question:

To what extent deep learning methods and techniques can be used to forecast the warehouse product demand for the future?

The primary goal of this study is to make advantage of the deep learning methods to predict the sales demand for a product for any given particular time in future. SARIMA and RNN models are going to be utilized to perform this research. The model will be trained using several hyperparameter combinations, and its performance will be evaluated using various evaluation criteria.

2 Related Work

To keep the inventory of any warehouse store updated might become an overwhelming task. As to meet the demand of a product depends upon various factors such as the location of the store, the time of the year and number of the users and so on. Demand forecasting has been playing a major role since the past in keeping track of those inventories. It is a best helping hand in managing the inventory.

2.1 Review Of Demand Forecasting Techniques Applied

The main purpose of any inventory management is to ultimately cut the cost incursion in material stocks. The paper Suesut et al. (2004) depicts the ways to prevent back logged demand and suggest ways to control the inventory level while lead-time. They also suggest how to manage the acquisition cost. In this paper researcher suggested ways to predict the optimal reorder level and reorder point using the order quantity. Which ultimately is helping in tracking the next cycle of the implied forecasting model. Computer integrated manufacturing (CIM) has been used to support the manufacture and design work whereas also helping the business management. The automated warehouse model uses programming logic controller (PLC) to control the operations. In this article they represented a way to incorporate factory automation systems by automating inventory control system and integrating it with computer networks, and information technology. Information technology and internet applications can help save time and money when it comes to communications among warehouses and end to end delivery. The paper showcases the important aspects like Just in Time (JIT) system, supply chain management system and logistics management could contribute a lot in changing the industries operational system.As a result, forecasting models are commonly used in precision marketing to understand and meet the wants and expectations of customers Seyedan and Mafakheri (2020).

Another way of predicting demand has been done in article Carbonneau et al. (2008), where author has used two data sets to conduct their research. The datasets that they used in the research, have been obtained from the simulated supply chain whereas the other dataset is gathered from actual Canadian Foundries orders. In this study they have applied support vector machines (SVM), regression models and some other advanced models as well. In later part they have compared the results using mean absolute error and standard deviation. They have explained end to end flow for a model of a long-distance supply chain with rising demand distortions and a cooperation barrier. They have also mentioned the limitation among the workflow as there is no explicit forecast information-sharing between parties in a supply chain. Author also highlighted that, if the forecasting accuracy can be improved, it is going to result in cheaper costs due to decreased inventory and higher customer satisfaction due to enhanced on-time delivery. They have also mentioned how distorted demand forecasting could lead into failure of entire workplans and manufacturing cycle.

Continuous reviewing and forecasting systems could help in automating the warehouses and controlling the inventory. The idea has been presented and a research has been conducted in article Suesut and Mongkhoin (2004). The order trends can be predicted using demand forecasting along with redefining the inventory reorder point (RP) using the automated warehouse system. One of the dominating factor for automated warehouse is computer integration manufacturing system (CIMS). CIMS is contributing in completing many tasks in multiple domains such as material cost accounting, quality control, process control, and also few production related tasks such as production planning, material need planning, work order generation and shipment planning. CIM is supposed to affect every department and its primary branches are management and planning functions, plant wide based data handling, automatic warehouses and stock control and Flexible manufacturing system (FMS). In this research they have designed a system model for inventory control by keeping economic order quantity and periodic review system into consideration.

In paper Hodžić et al. (2019) the importance of the fast delivery of the goods to the customers has been addressed. Modern warehouse management systems (WMS) are introduced to overcome the problem of improving and optimizing the warehouse processes. As per their understanding, to implement such kind of warehouse paperless management system can be used. They also described how mandatory it is to not clutter up the inventories with extra stock as the space is quite limited and expensive. Also, if the warehouses are overfilled with stuff, then it is only going to make delivery more delayed because of the consumption of the time in collecting the order. Long Short-Term Memory neural network (LSTM) has been used by the author to predict the demand of the product. The results obtained from the neural network has been compared with already existing basic mathematical model like median algorithm. After comparison they concluded that neural network showed lesser mathematical accuracy, which might change as per several more variables like times of weekends, discount, and holidays.

In research work Shantilal et al. (2019), ensemble approaches were used to do research on an alcohol warehouse in Bangladesh, where demand forecasting is a novel idea. Because Bangladesh is a small nation and demand forecasting is still emerging there, the results of this article can be used to compare models from other countries. There is a lot that can be applied there. Data from the previous week, month, and year has been compared to actual and expected product demand. Another research Islek and Oguducu (n.d.) provide a study that uses Linear Regression, SVM, Bayesian Networks, and a classifier ensemble to compare outcomes. Various forecasting approaches were combined together using neural networks in this study, and the findings of that model were then compared to the results of other separate models.

2.2 Supply Chain Management In Business Over Internet

SCM has its great impact on almost every industry type. One of the industry type has been showcased in research paper Tong and Dao-zhi (2010). Author has displayed supply chain effect on vendor managed inventory (VMI). In their research author has considered vendor managed inventory with fuzzy demands, which later on has been compared with a Retailer managed inventory (RMI) system. Diverse from the routine ponders, in this article author developed and fathoms a fuzzy VMI demonstrate. By the use of a triangular fuzzy number model has been created also the outside request, and the closed shape arrangements for VMI model are obtained.

Various problems which are faced by main distribution warehouses while predicting the demand for a product in any given time in future has been addressed by researchers in article İşlek and Öğüdücü (2015). It's difficult to anticipate how much of a product buyers will buy in the future since so many factors impact it. Among which primary once are including the size of the warehouse region, the number of customers, and the type of goods. For a chain of warehouse, it becomes more complex to predict the demand of customers. The author proposed a solution for this by using bipartite graph clustering, the suggested approach groups comparable warehouses based on their selling behaviour. Following that, a hybrid forecasting phase is used, which employs both a Bayesian Network and a moving average model machine learning method. By taking into account different warehouse, product, consumer demographic, and temporal variables, the suggested model depending on data mining methods is able to correctly anticipate product requests. They successfully clustered the data and got a error rate of 17% only for their third trial, which could be considered as an effective solution for this problem.

With increasing world population several countries are facing worst consequences. Problems like food security, poverty and hunger still remains constant and challenging in India and this problem has been addresses in article Sharma and Parhi (2017). They have tried to throw some light on how this big issue could be overtaken by using of big data warehousing which can be used in enhancing the accessibility of the food. The four core parameters that needed to keep in mind while talking about food security are availability of food, affordability of the food, stability of the food and accessibility of the food (Sharma and Parhi; 2017). The main cause for the system failure are the lack of information about the demand, storage, infrastructure facilities and market accessibility. With the help of advancement in big data structures, agriculture may be benefitted a lot and thus helping in improving productivity. Also better stock management and faster deliveries to the customers could be benefitted with this. Many challenges that are faced by big data applications are mentioned in the paper like cyber security, connectivity and decision support tools.

Forecasting uncertain demand is one of the most difficult tasks in Supply Chain Management (SCM) Nikolopoulos et al. (2016). Author has applied multiple forecasting methods like Croston's method, SBA, and TSB, as well as some more recent non-parametric advanced methods. However, none of them can detect or extrapolate data patterns, which is significant since these patterns arise regularly, driven by uncommon but nonetheless recurrent managerial activities. Author represented a way to overcome this situation by using Nearest Neighbours(NN) as it can pick up patterns in even shorter series. As per the research they have mentioned the main difficulty in Operations Management (OM) in the vast majority of supply chain and logistics systems, is estimating sporadic/intermittent demand as precisely as feasible. The study's findings imply that employing NN techniques can be beneficial to professionals. However, with huge datasets, these strategies should not be employed unsupervised. In place of that they can be utilized specifically in which a standard parametric strategy (SBA or TSB) is utilized by default and the NN is utilized as it were when dependable data verifying to the presence of designs is available.

SCM is an inventive and efficient perspective to tackle a variety of problems in development cooperation. The construction supply chain is relatively fractured and inefficient due to the unique construction management method and collaboration. In article Hu (2008) author represented an organized supply chain management model that can be used for the construction projects. The data stream within the development supply chain framework is modified to progress development collaboration. Also a compelling development collaboration method is given to quicken development venture administration developments. The bulk of development administration is project-based, and the development supply chain is fragmented, with isolated obligations from independent development members and separated development lifecycle stages. Effective development cooperation across all personnel and the development lifecycle is key to moving forward with development administration execution.

In research paper Xu (2006) a consolidated SCM model has been developed, influenced by market demand. Which is developed based on an evaluation of current integrated supply chain models, to solve the problem of growing supply chain costs in manufacturing organizations. An integrated supply chain framework, which incorporates the interior administration framework and the external supply chain framework of a manufacturing business, is presented to develop this integrated supply chain planning model. In addition, to depict the supply chain preparation, as well as a conceptual system of supply chain arranging coordinates. To achieve this integrated planning model has been applied. A 3 staged planning model, market demand prediction model and cost analysis model has been analysed to implement the best suited model. The relationship among production planning, logistic planning, purchase planning and sales planning has been established and considered by the author while implementing the model.

With the internationalization of the market and the advancement of innovation, the competition between businesses has shifted to supply chainsJiyun and Hu (2009). A symbiotic connection should exist between two members of a supply chain. However, they frequently clash with one another due to their disparate goals, which force everyone to behave in their own self-interest. The data is skewed in a tangible haggling when it comes to the manufacturer-retailer connection within the supply chain. The manufacturer is confronted with an ambiguous and arbitrary calculate external, but the retailer has all of the facts and its exterior calculate is already known. In this scenario, the manufacturer's premise is stable, but the retailer's has altered, causing supply and demand to be out of balance, resulting in a fight.

2.3 Demand Forecast For Hazard Prevention

Within the examination of crisis supplies on the premise of request characteristics, through the nearest neighbour strategy and combination of case-based thinking, fuzzy thinking and case-based thinking combines two sorts of crisis supplies for the foundation of request determining demonstrate and determining strategies utilized in conjunction with observational investigation, a case is, indicating out that the possibility of estimating strategies, crisis supplies for crisis needs by giving quick and doable forecast strategies. In research Rui (2010) two kinds of models are used for forecasting, in which one is using K-NN based reasoning methods to predict the demand of emergency supplies and the other model is using fuzzy reasoning under case-based reasoning methods.

It is very mandatory step for every organization to continuously track their everyday sales in order to forecast the sales of goods and services. If corrective and positive steps are undertaken for the predicted data, then it could benefit any organization in increasing their sales, so that demand supply ratio could be maintained. In research Katkar et al. (2015), sales has been predicted using Naïve Bayesian classifier. They have used a dataset which has been collected from different shops located in multiple cities. As the collected data is too diverse, it might be considered as a good dataset collection type to be consider for future research as well. Fact tables has been created by using primary keys of each dataset. Star schema and fact constellation schema has been used to achieve this task. As the data for a warehouse store could reach upto a huge amount and thus to train such data could become an overwhelming task. Author used sales order table which were extracted from OLTP system.

Any uncertain situation can arise anytime, a good example for this is COVID-19 virus. Which affected market very badly. In research paper Srivatsa Srinivas and Marathe (2021) this issue has been addressed and the measures which could be taken to deal with these kinds of issues has been explained. For warehouses and stores they have suggested to replace the rigid warehouses with mobile warehouse. A mobile warehouse could be a big vehicle which is dedicated for a particular region, which only carries the stock of different products as per the demand forecasted for that region. Furthermore, since the conveyance time is anticipated to be brief, customers can helpfully collect the arrange on request, and costs related with the nonappearance of buyers are maintained a strategic distance from.

Another research paper Kim (2020) has discussed the same issue as mentioned above related to global pandemic. In the interim, how COVID-19 has impacted buyers and the utilization culture has gotten generally constrained consideration. Directors regularly take a wait-and-see approach on the affect of COVID-19 on deals. In this paper, the widespread as a quickening agent of the auxiliary alter in utilization and the computerized change within the commercial centre has been addressed. Supervisors might adjust to the advanced change within the advertise to recuperate or indeed develop encourage the deals after COVID-19. In a recent survey it has been noticed that for online shopping consumer's behaviour can be changed not just because of network effects but also it can be changed due to pandemic effects. As per a recent online survey in USA 37% of people considered to shift to online purchasing as an adverse effect of COVID-19.

2.4 Summary of Literature Review

This research effort has been influenced by the studies and research stated above. In second subsection of the review, the importance of supply chain management has been showcased and its impact on different industries. Whereas, in first section different methodologies and research that have been done on demand forecasting has been highlighted. Few deep learning methods that have been applied also displayed some positive result but because of several limitations the research still could be improved. Because of lack of data and lack of technology there is still a huge area of modification in warehousing sector, as this domain is very vivid and huge.

3 Methodology

The methods that have been applied in the past are working at good efficiency. However, there is still a chance for improvement and any advancement in feature could led to a great help. The most common problem with demand forecasting is the availability of the data and location. Another issue with basic prediction model is the overfitting and thus lead to wrong interpretation of the results. To overcome these two challenges, data has been collected which includes historical product demand for a worldwide manufacturing corporation. Four warehouse data has been used which are placed at different locations and ship goods as per their allocated regions. In terms of researching further and applying best fitted model, Recurrent Neural Network (RNN) has been used.

To carry out this research, Cross-industry standard data mining process (CRISP-DM) methodology has been used. CRIP-DM has six important aspects which might be considered while applying which are displayed in Figure 2.

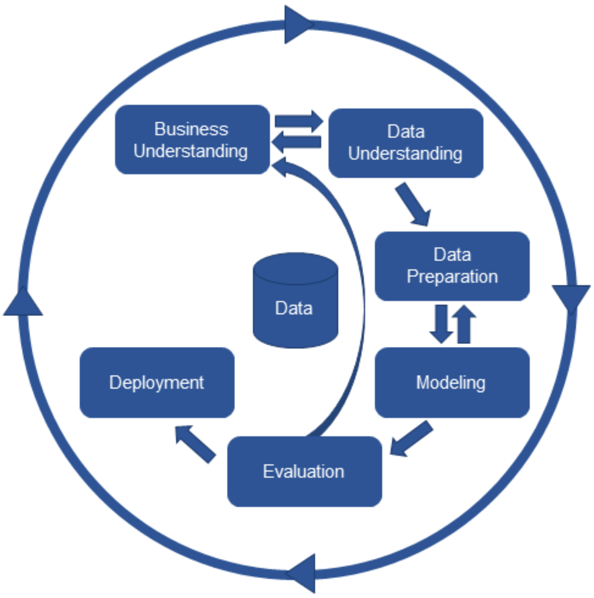

Figure 2: CRISP-DM Model

3.1 Business Understanding

Deep learning methods could contribute a lot in e-commerce sector as E-Commerce is continuously emerging and growing sector. New technologies and advanced neural networks could benefit in more precise prediction of products demand, so that warehouses can be prepared in advance for upcoming events. Already existing methods and implementation could be used further for applying on advanced deep learning methods. If more precise predictions could be made, then it is going to help in good business growth as well as more customer satisfaction. Usually, it takes approximately over a month to transport the products through ships. Using effective product demand forecasting techniques, the time consumption can be reduced for the next cycle of the supply. The main aim of this research is to find an efficient way to predict and forecast the product demand for a particular time in the future more accurately.

3.2 Data Understanding

The dataset that is being used for this research has been downloaded from public repository Kaggle. The data collection includes historical product demand for a worldwide manufacturing corporation. The data has been divided into thousands of categories and multiple product categories. There are four central stockrooms to transport items inside the locale it is dependable for. Since the items are made totally different areas all over the world. The used dataset has string, datetime and integer values and thus it can be used straight away for the research. The proposed model has been implemented, trained, tested, and verified on this dataset only.

Firstly, for the dataset after performing all the cleaning operations, box plot has been made for order demand. Another box plot with log transformation has also been made. From Figure 3 distribution of the data could be noticed.

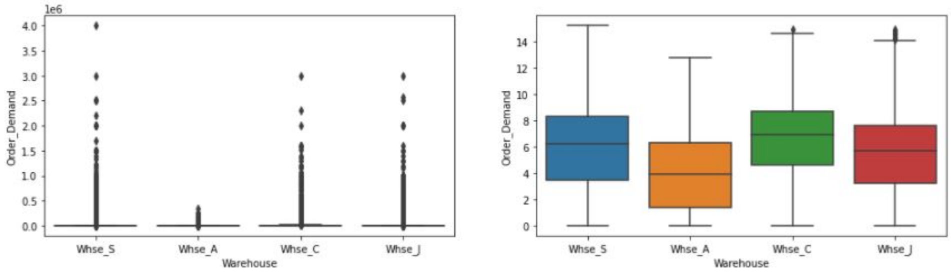

Figure 3: Distribution of Warehouse data

Once done with pre-processing of the data, with the help of with decomposition analysis, data has been broken down to analyse seasonality, trend and residuals. As could be seen in Figure 4, the sales are comparatively low in the start of the season, and demand peaks in the fourth quarter of each year.

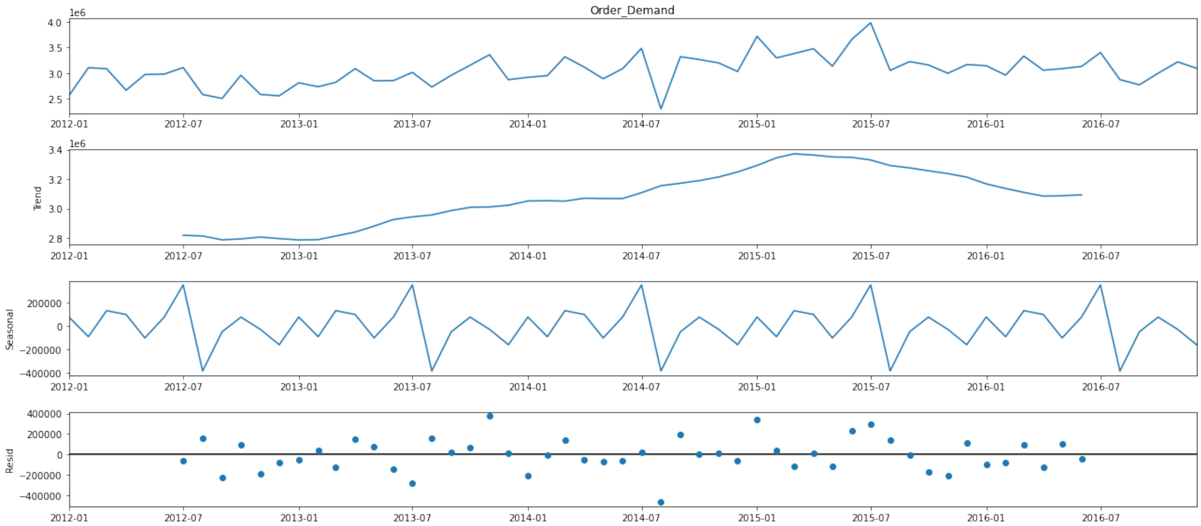

Figure 4: Decomposition of Time Series Data

3.3 Data Preparation

The downloaded dataset suppresses a total of five columns which include Warehouse, Date, Product Category, Product_Code, and Order Demand. The column data is in string, datetime and integer format. The dataset is neither huge nor small, it sizes 49Mb which might be considered as good dataset for this study. As two different models are being applied and compared, for applying SARIMA data can be used straight away. However, for applying RNN data needs to be manipulated and datatypes need conversion to be fed to neural network.

3.4 Data Pre-processing

The chosen dataset has 10,48,576 rows and 5 columns which are provided for train data and test data. There are several entries which has no data related to date and as this model is going to be predicting by keeping date as important factor and thus those null value data has been filtered out. There was a total of 11,239 values which did not have any date assigned. However, most of these values belonged to year 2011 and 2012 which might be considered as irrelevant to our research and thus decided to be dropped. So, data has been capped year ranging from 2012 till 2017. For the target variable that is order demand the format needed to be changed from string to integer, also few values consisted of brackets, so for the research they were also to be removed. To apply neural network date column had to split into 3 new columns date, month, and year.

3.4.1 Encoding the Data

After that the categorical variables have been converted to continuous variables using label encoder. LabelEncoder method from sklearn library has been used to convert these values.

3.4.2 Splitting the Data

To test data 5000 records were kept, also to verify the result 5000 more records were kept on hold. The total train test and hold split used for research is 1038575, 5000, 5000 respectively.

3.4.3 Preparation for 3-Dimensional Sequential Data

Preparing a 3D input data is big challenge in itself. Sequential models have 3 dimension (test estimate, time steps, highlights). Rather than attempting to make a 3D cluster, in this paper TimeSeriesGenerator class has been utilized which brings a few other preferences like setting the batch size.

3.5 Modeling and Evaluation

The models that have been applied for the research are described in this section. SARIMA and Recurrent neural network has been applied to conduct this research.

SARIMA is an abbreviation used for Seasonal Auto Regressive Integrated Moving Average, which can reflect a variety of common temporal features in time series dataWaheeb et al. (2019). Seasonality is the additional factor in SARIMA which is added to normal ARIMA model. In this model, autoregression is utilized to contribute to the dependent connection between an observation and a number of delayed observations. Integration has an impact on the usage of raw observation differencing to keep the time series constant.A residual error from a moving average model applied to lagged data is utilized in this model to illustrate the reliance between an observation.

In RNN, recurrent indicates that the present time step's outcome becomes the next time step's input. The model analyzes not just the current input, but also what it knows about the previous elements at each element of the sequence. As displayed in Figure 5

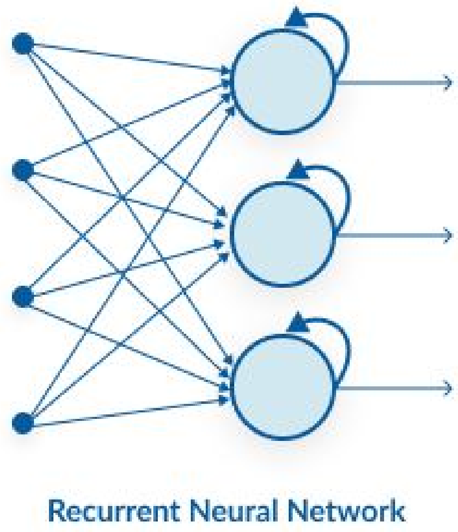

Figure 5: Recurrent Neural Network

that the output of one layer works as the input for second layer, thus this neural network keeps on learning while training phase itself.

To evaluate the models different approaches have been used for both. To evaluate SARIMA model mean squared error(MSE) and root mean squared error(RMSE) have been calculated to get better understanding of the model. Whereas for the RNN model mean absolute error(MAE) and correctness of the model has been calculated for both training and testing datasets.

4 Design Specification

In this section SARIMA and RNN functioning has been described, as they are the model which has been used to solve the problem.

4.1 SARIMA(p, q, d)(P, Q, D)m

The model is using statistics model SARIMAX for implementing SARIMA model. Series decomposition has been carried out to check seasonality and trend of the data, also to check the residuals of the data. While implementing SARIMA below mentioned parameters are to be kept into consideration:

- S: Seasonality

  Seasonality refers to recurring patterns, such as increases or decreases in activity due to a variety of circumstances.

- AR: AutoRegression

  AR stands for auto regressive, which signifies to forecast time series values based on previous periods.

- I: Integrated

  This is an upward or downward tendency which appears in the trend of the data. Differencing can be used to get the data free from this tendency.

- MA: Moving Average

  Moving average is responsible for passing the error values from the past cycle to the next iteration of the network.

- (p, q, d):

  This component describes hyper-parameters values. AR value is passed as parameter p in this model which is the total number of lag order. MA value is parameterized as q and represents the order of moving average. Integrated value can be passed by setting up value of d, which represents the total count when the raw observations have been differentiated.

- (P, Q, D):

  The second component for the SARIMA is (P, D, Q) which represents the analogs of (p, q, d), whereas the season component is separated form it.

- m

  m signifies the factor of seasonality. It basically defines the total number of time steps which are present for one period of a season.

4.2 RNN

RNNs are a sort of neural network that remembers what it has processed previously and can thus learn from past iterations during training. As by using a feed-forward network, it cannot work on the previous set of the data, this challenge can be overcome by using RNN.

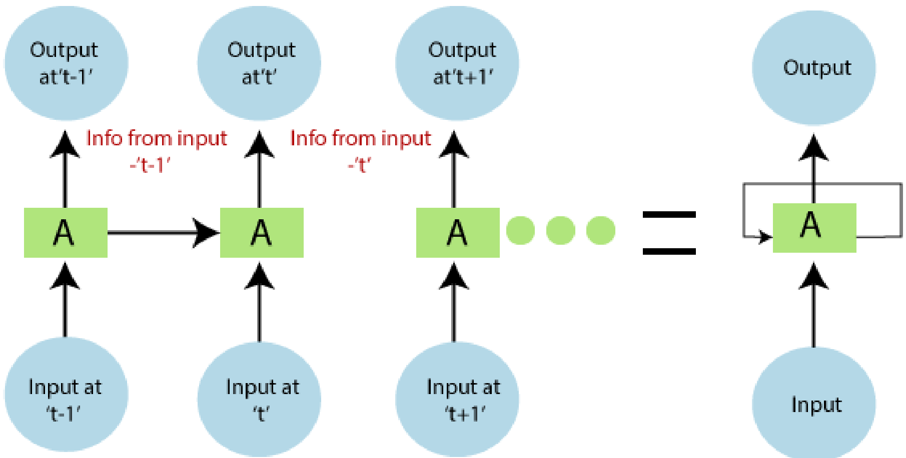

Figure 6: Recurrent Neural Network

As it can be interpreted from Figure 6, knowledge from the previous input 't-1' is being provided along with input at the next step to provide the output at stage 't'.

When information is forwarded to next layer, it gets connected with an edge(Connection line). Edge consist of the total net input and sums up all the inputs. A transfer function is used to calculate the output, on which it is dependent that what amount of that node is going to contribute in next layer's input. To conduct this research 'Linear' and 'ReLU' functions are used.

The passage of the whole training dataset through the network, comprising of one forward and one backward pass, is referred to as an epoch. The number of epochs used to train the RNN will decide the trade-off between training time and accuracy. In this research different numbers of epochs were used to obtain the best result.

5 Implementation

This chapter covers the complete research process. It goes over the setup procedure, data processing steps, model execution, and tools used.

5.1 Environment Setup

Table 1: System Configuration

	System Configuration
Operating System	Windows 10 Home Single Language 64-bit
Memory	16.0 GB RAM
CPU	Intel(R) Core(TM) i7-8550U CPU @ 1.80GHz
Cores	8
GPU	Intel(R) UHD Graphics 620

Programming language: Python 3.8.12

IDE: Jupyter Notebook

5.2 Data Processing

The dataset utilized in this study was obtained from the Kaggle public repository [1]. Thousands of categories and product categories have been used to organize the data. String, datetime, and integer values are all included in the collection. The selected dataset comprises 10,48,576 rows and 5 columns for train and test data, respectively. There are some items that have no date data, and because our model will be forecasting using date as a key component, those null value data have been filtered out. The majority of these figures pertained to the years 2011 and 2012, thus they were removed. The data has been blocked for the years 2012 to 2017. The format of the goal variable, order demand, had to be changed from string to integer, and a few numbers had brackets, thus they had to be eliminated for the study. The date column has to be separated into three new columns date, month, and year to apply neural network. The label encoder was used to transform the categorical data to continuous variables. 5000 records were held on hold to test the

[1] https://www.kaggle.com/felixzhao/productdemandforecasting/

data, and another 5000 records were kept on hold to validate the results. Data has been transformed into 3-Dimensional sequential data using TimeSeriesGenerator class.

5.3 Model Training

To implement SARIMA model, all the parameters values (p,q,d,P,Q,D) were passed by iterating through the loop. So that the best fitted model could be found out. Then Akaike information criterion (AIC)[2] values for each iteration for fitting model by SARIMAX()[3] function has been collected. The best result was obtained when variable value is set to ARIMA(1, 1, 1)x(1, 1, 0, 12)12. After feeding these values to the model, diagnostics plots have been plotted and values are predicted. After that MSE and RMSE values have been calculated as per the outcomes.

To train RNN model, node value is taken as 4, activation function is taken first as relu and tried to be applied as linear for RNN and output layer both. Adam is the optimizer. The mean squared error is the loss function. The rate of learning is taken as 0.0001. Epoch values have been changed and analyzed upon values 50, 100 and 200. MAE values and correctness values have been measured after receiving the output.

Different combinations of variables that have been considered during research are as explained in table 2:

Table 2: Variables Configuration

	Variables Configuration	
Epoch: 50	Activation Function: Linear	Hold Data: 500
Epoch: 500	Activation Function: Linear	Hold Data: 5000
Epoch: 50	Activation Function: ReLU	Hold Data: 500
Epoch: 500	Activation Function: ReLU	Hold Data: 5000

6 Evaluation

The outcomes of the applied model are described in details in this section.

6.1 SARIMA

All of the parameter values (p,q,d,P,Q,D) were provided by iterating through the loop to implement the SARIMA model. So that the most appropriate model may be identified. Figure 7 is the snapshot for the outcome of the loop. The Akaike information criterion (AIC) value is used to identify the best outcome, which is a measure of a statistical model's relative quality for a given set of data. The AIC is a metric that quantifies how well a model fits the data while taking into consideration the model's total complexity. The smaller the AIC value, Model fits similar but using lesser features, and thus for this research the AIC value chosen is 960.51 which is received for ARIMA(1, 1, 1)x(1, 1, 0, 12)12. Thus this set of ARMIA variables has been used.

[2]https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/AIC

[3]https://www.statsmodels.org/stable/examples/notebooks/generated/statespace_sarimax_stata.html

```
SARIMA(0, 0, 0)x(0, 0, 0, 12)12 - AIC:1932.23655778549
SARIMA(0, 0, 0)x(0, 0, 1, 12)12 - AIC:1512.9275832124356
SARIMA(0, 0, 0)x(0, 1, 0, 12)12 - AIC:1338.8201294951011
SARIMA(0, 0, 0)x(0, 1, 1, 12)12 - AIC:3134.0602952352074
SARIMA(0, 0, 0)x(1, 0, 0, 12)12 - AIC:1366.5117513512635
SARIMA(0, 0, 0)x(1, 0, 1, 12)12 - AIC:1340.8450308457732
SARIMA(0, 0, 0)x(1, 1, 0, 12)12 - AIC:1023.6756022859483
SARIMA(0, 0, 0)x(1, 1, 1, 12)12 - AIC:3025.077015756434
SARIMA(0, 0, 1)x(0, 0, 0, 12)12 - AIC:1862.087487804522
SARIMA(0, 0, 1)x(0, 0, 1, 12)12 - AIC:1471.183803270069
SARIMA(0, 0, 1)x(0, 1, 0, 12)12 - AIC:1305.328981334548
SARIMA(0, 0, 1)x(0, 1, 1, 12)12 - AIC:3048.8520947613288
SARIMA(0, 0, 1)x(1, 0, 0, 12)12 - AIC:1529.100572165093
SARIMA(0, 0, 1)x(1, 0, 1, 12)12 - AIC:1467.6395590277816
SARIMA(0, 0, 1)x(1, 1, 0, 12)12 - AIC:1020.634762975912
SARIMA(0, 0, 1)x(1, 1, 1, 12)12 - AIC:2969.939154212284
SARIMA(0, 1, 0)x(0, 0, 0, 12)12 - AIC:1648.7378898187837
SARIMA(0, 1, 0)x(0, 0, 1, 12)12 - AIC:1309.865329210198
SARIMA(0, 1, 0)x(0, 1, 0, 12)12 - AIC:1318.7588141990293
SARIMA(0, 1, 0)x(0, 1, 1, 12)12 - AIC:3112.67993693746
SARIMA(0, 1, 0)x(1, 0, 0, 12)12 - AIC:1331.924340756696
SARIMA(0, 1, 0)x(1, 0, 1, 12)12 - AIC:1315.7243994326855
SARIMA(0, 1, 0)x(1, 1, 0, 12)12 - AIC:998.700939745945
SARIMA(0, 1, 0)x(1, 1, 1, 12)12 - AIC:3093.075165031324
SARIMA(0, 1, 1)x(0, 0, 0, 12)12 - AIC:1590.336917752371
SARIMA(0, 1, 1)x(0, 0, 1, 12)12 - AIC:1258.4897692522468
SARIMA(0, 1, 1)x(0, 1, 0, 12)12 - AIC:1272.6101180952066
SARIMA(0, 1, 1)x(0, 1, 1, 12)12 - AIC:3020.7246585859752
```

Figure 7: SARIMA Implementation

After applying SARIMA model the output has been evaluated first using diagnostics plots, refer to Figure 8 for details. As for good model, the distribution of residuals should be normal. In the histogram plot it is supposed that KDE line should follow closely the standard notation mean line. The ordered distribution of residuals matches the linear trend of the samples in the qq plot. The correlogram (i.e. Autocorrelation) diagram shows the time series residuals has low correlation with its own logged version.

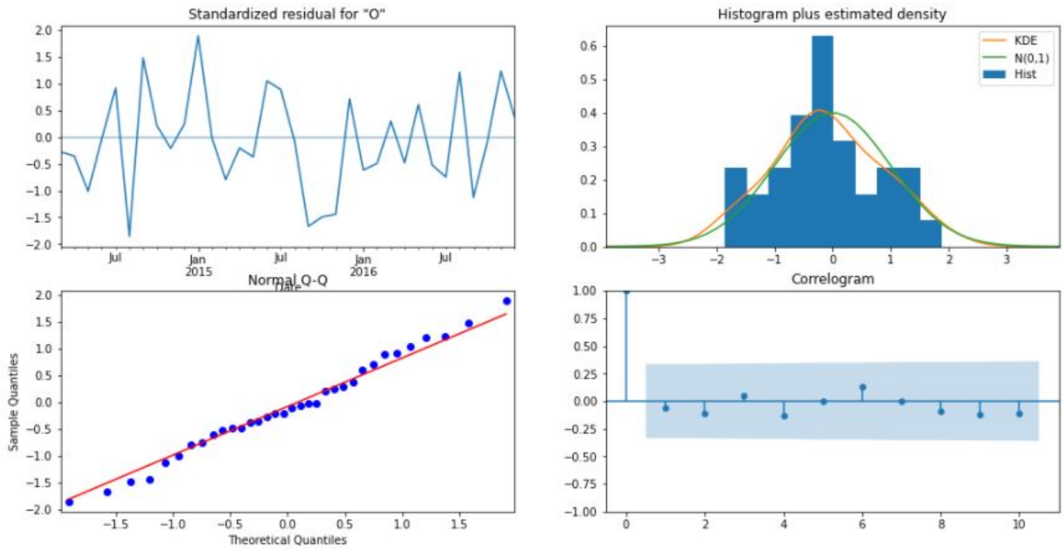

Figure 8: Model Diagnostics

After implementing all above, finally the model has been evaluated using root mean square error values. The MSE and RMSE for this model is received as 55459437820.56 and 235498.28 respectively.

```
: #Getting the mean squared error (average error of forecasts).
y_forecasted = pred.predicted_mean
y_truth = y['2016-01-01':]
mse = ((y_forecasted - y_truth) ** 2).mean()
print('MSE {}'.format(round(mse, 2)))

MSE 55459437820.56

: print('RMSE: {}'.format(round(np.sqrt(mse), 2)))

RMSE: 235498.28
```

Figure 9: RMSE for SARIMA

6.2 RNN

To train the RNN model, the activation function was set to relu and linear for both the RNN and output layers. Epoch values were altered and assessed at 50, 100, and 200. After obtaining the output, MAE and accuracy values were calculated. The following are some of the different outputs that are obtained after providing different combinations of variables:

6.2.1 Case 1

When Epoch has been taken as 50 with activation function as linear and hold data is set to 500 then model outcome is observed as below:

```
mean:  5401448.065173117
mae: 3824812.2773238607
mae/mean ratio:  70.81086832964439 %
correctness:  29.189131670355607 %
```

Figure 10: RNN Output 1

6.2.2 Case 2

When Epoch has been taken as 500 with activation function as linear and hold data is set to 5000 then model outcome is observed as below:

```
mean:  5399364.85453817
mae: 3070834.150571028
mae/mean ratio:  56.87398857645247 %
correctness:  43.12601142354753 %
```

Figure 11: RNN Output 2

6.2.3 Case 3

When Epoch has been taken as 50 with activation function as ReLU and hold data is set to 500 then model outcome is observed as below:

```
mean:  5399364.85453817
mae: 3585639.676204473
mae/mean ratio:  66.4085456864568 %
correctness:  33.591454313543196 %
```

Figure 12: RNN Output 3

6.2.4 Case 4

When Epoch has been taken as 500 with activation function as ReLU and hold data is set to 5000 then model outcome is observed as below:

```
mean:  6389941.9
mae: 4850469.7671875
mae/mean ratio:  75.90788528433256 %
correctness:  24.092114715667435 %
```

Figure 13: RNN Output 4

As the maximum correctness has been reached as 43.12% for epoch as 500 and when activation function is taken as linear, hence the rest of the study has been made for these set of variables only.

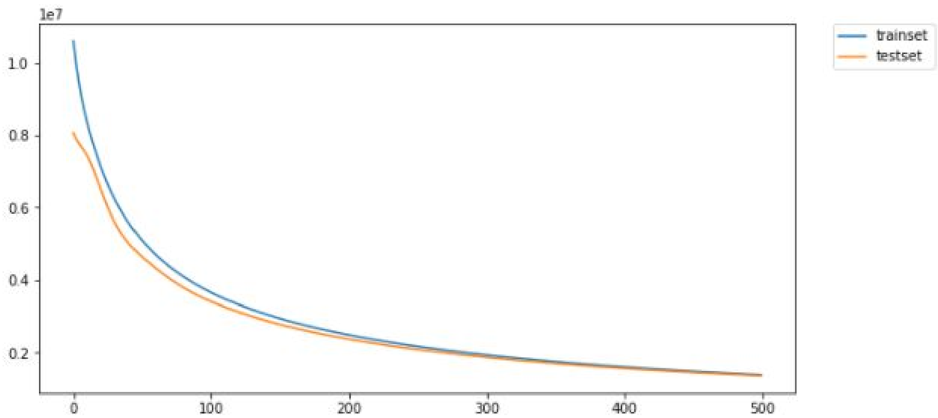

Figure 14: Loss Curve

As shown in Figure 14, the model sowed a high learning rate. Also, the actual vs prediction data graph is displayed in Figure 15.

```
plt.figure(figsize=(15,10))
plt.plot(df_result['Actual'], color='blue')
plt.plot(df_result['Prediction'], color='red')
plt.show()
```

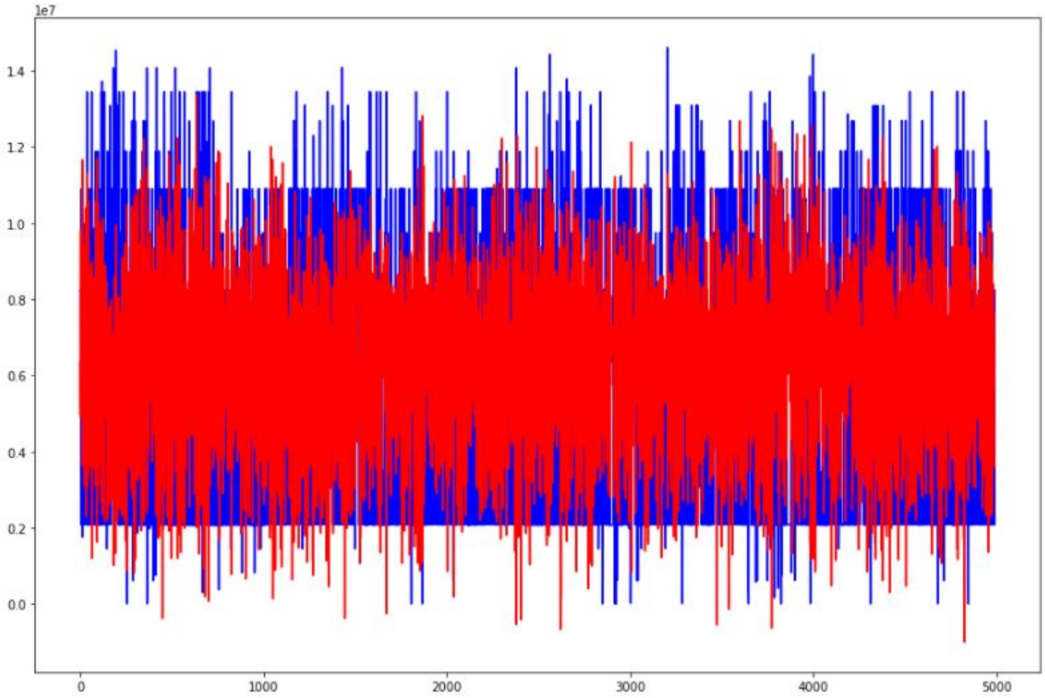

Figure 15: Recurrent Neural Network

6.3 Discussion

After receiving all the results from both the models, it can be established that among RNN there is still a lot of improvement that is needed to improve the correctness of the model. Whereas SARIMA implementation has been implemented quite appropriately and also the residual plot is distribute evenly along the standard normal distribution. AIC value that is received for set of variable ARIMA(1, 1, 1)x(1, 1,0, 12)12 is also very moderate. Thus giving the RMSE value of .23 which is considered to be good for fitting a model.

In RNN model, for the selected set of variables, the learning curve seems to be quite good, as it is intended more towards the origin. However, correctness value is not satisfying. Even though the correctness seems to be increasing across the spectrum. Staring from bare as minimum of 24%, with different variables it reaches till 43%, whereas keeping the MAE as 75%.

7 Conclusion and Future Work

Despite the fact that there is so much work that could be done in the warehouse business in context of Supply chain management and inventory management. This research done should assist in providing a viable solution for implementing a good model and make a better understanding in product demand forecasting. In the proposed study, many machine learning and performance evaluation methodologies will be applied. Using an advanced neural network for demand forecasting is a step toward improving the goods warehouse's current situation. If warehouses properly handle merchandise in accordance with current demand, customers will have a better experience, and corporations will benefit more. The suggested project will benefit all players in the warehousing business.

ACKNOWLEDGEMENT

Prof. Hicham Rifai deserves special mention for his unwavering support, recommendations, and surveys. I owe him a debt of gratitude for always encouraging and leading me to enhance my research. I would also like to encourage the efforts made by my classmates and friends for their consistent support.

References

Carbonneau, R., Laframboise, K. and Vahidov, R. (2008). Application of machine learning techniques for supply chain demand forecasting, *European Journal of Operational Research* **184**(3): 1140–1154.
URL: *https://www.sciencedirect.com/science/article/pii/S0377221706012057*

Hodžić, K., Hasić, H., Cogo, E. and Jurić, (2019). Warehouse demand forecasting based on long short-term memory neural networks, *2019 XXVII International Conference on Information, Communication and Automation Technologies (ICAT)*, pp. 1–6.

Hu, W. (2008). Improving construction collaboration performance through supply chain control and management, *2008 International Conference on Information Management, Innovation Management and Industrial Engineering*, Vol. 1, pp. 58–61.

Islek, I. and Oguducu, S. G. (n.d.). A decision support system for demand forecasting based on classifier ensemble.

Jiyun, J. and Hu, C. (2009). Supply and demand conflict in supply chain based on principal-agent theory, *2009 Second International Conference on Intelligent Computation Technology and Automation*, Vol. 4, pp. 11–14.

Katkar, V., Gangopadhyay, S. P., Rathod, S. and Shetty, A. (2015). Sales forecasting using data warehouse and naïve bayesian classifier, *2015 International Conference on Pervasive Computing (ICPC)*, pp. 1–6.

Kim, R. Y. (2020). The impact of covid-19 on consumers: Preparing for digital sales, *IEEE Engineering Management Review* **48**(3): 212–218.

Nikolopoulos, K. I., Babai, M. Z. and Bozos, K. (2016). Forecasting supply chain sporadic demand with nearest neighbor approaches, *International Journal of Production Economics* **177**: 139–148.
URL: *https://www.sciencedirect.com/science/article/pii/S0925527316300391*

Rui, L. (2010). Research on emergency supplies demand forecasting model, *2010 The 2nd International Conference on Industrial Mechatronics and Automation*, Vol. 1, pp. 244–247.

Seyedan, M. and Mafakheri, F. (2020). Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities, *Journal of Big Data 2020 7:1* **7**: 1–22.
URL: *https://journalofbigdata.springeropen.com/articles/10.1186/s40537-020-00329-2*

Shantilal, S. P., Gangil, M. and Student, M. (2019). Warehouse sales forecasting using ensemble techniques, *Research Journal of Engineering Technology and Management* **12**: 2582–2610.
URL: *www.rjetm.in/*

Sharma, R. and Parhi, S. (2017). A review on use of big data in warehousing to enhance accessibility of food, *2017 2nd International Conference on Communication and Electronics Systems (ICCES)*, pp. 663–667.

Srivatsa Srinivas, S. and Marathe, R. R. (2021). Moving towards "mobile warehouse": Last-mile logistics during covid-19 and beyond, *Transportation Research Interdisciplinary Perspectives* **10**: 100339.
URL: *https://www.sciencedirect.com/science/article/pii/S2590198221000464*

Suesut, T., Gulphanich, S., Nilas, P., Roengruen, P. and Tirasesth, K. (2004). Demand forecasting approach inventory control for warehouse automation, *2004 IEEE Region 10 Conference TENCON 2004.*, Vol. B, pp. 438–441 Vol. 2.

Suesut, T. and Mongkhoin, B. (2004). Demand forecasting approach inventory control for cims, *ICARCV 2004 8th Control, Automation, Robotics and Vision Conference, 2004.*, Vol. 3, pp. 1869–1873 Vol. 3.

Tong, A. and Dao-zhi, Z. (2010). A supply chain model of vendor managed inventory with fuzzy demand, *2010 International Conference on System Science, Engineering Design and Manufacturing Informatization*, Vol. 2, pp. 15–18.

Waheeb, W., Ghazali, R. and Shah, H. (2019). Nonlinear autoregressive moving-average (narma) time series forecasting using neural networks, *2019 International Conference on Computer and Information Sciences (ICCIS)*, pp. 1–5.

Xu, L. (2006). An integrated supply chain planning model for manufacture firm based on logistics chain, *2006 International Conference on Service Systems and Service Management*, Vol. 1, pp. 589–593.

İşlek, and Öğüdücü, G. (2015). A retail demand forecasting model based on data mining techniques, *2015 IEEE 24th International Symposium on Industrial Electronics (ISIE)*, pp. 55–60.