The Deep Learning Approach In Demand Forecasting In The Supply Chain Process Of Ice Creams At The Grocery Store Firm

Tran Luu Thi[1], Tran Tien Huy[1], and Ngo Thi Trang Thuy[1], Phan Tien Thanh[1], and Huu Thuan Thang Nguyen[1], Nguyễn Thị Dung[1], Lê Thị Long Châu[2], Nguyen Nhat Minh[3], Nguyen Ho Thanh Dat[1], Hoang Van Hai[1*]

[1] University of Economics - The University of Danang, Danang, Vietnam

[2] Danang Vocational Tourism College, Danang, Vietnam

[3] RMIT University, 702 Nguyen Van Linh Blvd, District 7, Ho Chi Minh City, Vietnam

[*]Corresponding author: haihv@due.edu.vn

Abstract

Research purpose:

This study investigates the application of a Deep Feedforward Network (DFN) for demand forecasting and customer willingness-to-pay predictions in the retail sector.

Research design, approach, and method:

Using real-world data from VinMart and other retail stores, the DFN was tested for its ability to predict order quantities across multiple locations.

Main findings:

While the model delivered promising results, training separate networks for individual stores proved more effective than using a single network for all stores, due to varying input significance levels. Data limitations, such as the lack of extensive historical records, affected accuracy. Additionally, the DFN was used to predict customer willingness-to-pay, with inputs gathered through quantitative research. While the model showed potential, its success is highly dependent on data quality.

.

Practical/managerial implications:

Future research should explore alternative deep learning architectures and incorporate more diverse variables to enhance accuracy. This study underscores the importance of robust data for optimizing supply chain decisions in retail..

Keywords: Deep learning, Supply chain, Ice-cream industry, Inventory level control.

D. Nguyen Van et al. (eds.), *Proceedings of the International Conference on Emerging Challenges: Sustainable Strategies in the Data-driven Economy (ICECH 2024)*, Advances in Economics, Business and Management Research 320,
https://doi.org/10.2991/978-94-6463-694-9_23

2. INTRODUCTION

In the context of Artificial Intelligence (AI) which is growing stronger and easier to replace manual jobs to save labour, Vietnam is one of the countries determined to focus on development. AI technology, which is considered a spearhead needs to be researched and forecasted to be the most breakthrough technology industry in the upcoming years [1]. The supply chain is a complex system that is affected by many elements, from manufacturing sites and warehouses to transportation, inventory management, and order fulfilment. An optimized supply chain can fulfil the needs of customers, and reduce the cost of stock and delivery. Many companies and enterprises are working on ways to optimize the chain and in particular, demand forecasting techniques are required, especially in the market with fierce competitive and flexible demands. However, not every demand forecasting method gives an accurate value and some of them may not be capable of dealing with huge data. According to McKinsey (2017), “AI-enhanced supply chain management greatly improves forecasting accuracy while simultaneously increasing granularity and optimizing stock replenishment. Reductions between 20 and 50% in forecasting errors are feasible. Lost sales due to products not being available can be reduced by up to 65% and inventory reductions of 20 to 50% are achievable” [2]. Because of this, Deep Learning Approach is proposed to improve forecasting accuracy and therefore, help companies to manage their supply chain in a better way.
The accuracy of deep learning is calculated by comparing the predicted value with the real value or using the MAPE function [3]. As the accuracy of an approach is appropriate and contains fewer errors overall, the approach can improve the outcome of the demand forecasting process. The accuracy of deep learning approaches positively affects the outcome of demand forecasting in the supply chain.

The deep Learning approach to demand forecasting is basically data forecasting based on historical data. Because of that, bad data, inappropriate collecting methods, and wrong variable selection can lead to inaccurate output or conclusions. However, defining the quality of data is challenging so in order to overcome this issue, improving the performance of data collecting is recommended, or testing the model widely with a wide range of data. Furthermore, it is a possibility that some unexpected variables occur in the process of the system. Those variables are considered as if they never happened in the past or occur too randomly. Because of that, it is almost impossible to collect data to fit those variables and therefore, the outcome of the model when this happens can be irrelevant. To sum up, the performance of historical data collecting methods has an impact on the outcome of demand forecasting in the circumstance with no unexpected variable occurring
.

3. LITERATURE REVIEW

3.1. Sustainable manufacturing (SM) practices

1.1 Deep learning

According to Kwang Gi Kim (2016), Deep Learning (DL) is a form of machine learning that enables computers to learn from experience and understand the world in terms of a hierarchy of concepts [4]. Because the computer gathers knowledge from experience, there is no need for a human computer operator formally to specify all of the knowledge needed by the computer. The hierarchy of concepts allows the computer to learn complicated concepts by building them out of simpler ones; a graph of these hierarchies would be many layers deep. DL has already proven useful in many software disciplines, including computer vision, speech and audio processing, natural language processing, robotics, bioinformatics and chemistry, video games, search engines, online advertising, and finance. According to Artemm Oppermann (2019), deep learning is a subset of machine learning in which multilayered neural networks learn from vast amounts of data [5]. The neural network of deep learning works like the human brain and it receives information, then analyzes the information after coming up with several outputs. Deep Learning is able to learn from vast amounts of data and it is useful for our model which allows DL to learn and predict data based on what they learned [5].
In this study, we use a feedforward neural network as an algorithm for deep learning. This network is considered useful in classification and forecasting. According to Tushar Gupta (2020), a feed-forward neural network is an algorithm that contains several units, organized into 3 different layers: Input layer, hidden layer, and output layer [6]. Data enters the network at the input layers and then is processed by nodes in a hidden layer. Nodes in a feedforward network are basically functions, that can be linear or nonlinear. The functions are adjusted by the network by changing the weight or bias of each node. Those changes in weight will lead to a change in the output value. There are no feedback connections in this network, the data goes through the input layer to the output layer, layer by layer, and the outputs are fed back into itself. Feedforward neural networks can be used for forecasting data. To do this, the system must be trained with a training sample, including training input and training output.

1.2 Supply Chain demand

Therefore, these demands will be affected by different factors, so specific standards are needed to define parameters. The term "demand" has different meanings for different customers [7]. The definition of “need” also has different meanings for different research papers. As a result, various factors will influence these demands, necessitating the development of precise criteria to define parameters. For various consumers, the word "demand" has different meanings [7]. The term "need" has different connotations in different research articles. This is an advertisement. Increasing demand –or demand-and Supply Chain Management (DSCM) can be clarified by the understanding that there is only a combination [8]. New businesses may be able to adapt to consumer needs more quickly and effectively than older businesses. By minimizing delivery times, improving service quality, and lowering prices, collaboration between suppliers, distributors, and retailers may help increase the number of satisfied customers. White (1997) and Pearce (1986) described "the need for quality" as the need for goods and services to be provided by capital to pay for them [9], [10]. People's desire for something will not be accompanied by a willingness to pay for it.

1.3 ANN Model

ANN, Artificial Neural Network, is inspired by the human brain and it is generated by a hidden layer, which is between the input layer and output layer. The ANN can do many tasks in many different fields and it is able to handle a huge number of data. Because of this, ANN has become a significant technology currently while Big Data is a trend in many industries.

Forecasting is a major use of ANN and it is widely used as it provides an efficient tool for many businesses. Weather forecasting, energy consumption forecasting, trend forecasting, etc, can be done by ANN as long as the system contains enough data. ANNs are data-driven self-adaptive methods, that can learn from examples, known as training data, and they capture functional relationships between data [11]. In a forecasting model using ANN, after the model learns the data given and captures the relationships, the model can then predict future data. ANNs are nonlinear and in real-world systems, especially in any forecasting model, nonlinear function is more likely to occur. ANNs are capable of performing nonlinear functions without finding the pattern between input variables and output variables at first. Because of this, ANN is considered flexible and accurate at forecasting tasks.

According to Tushar Gupta (2020), a Feedforward neural network is an algorithm that contains several units, organized into 3 different layers: Input layer, hidden layer, and output layer [6]. Data enters the network at the input layers and then is processed by nodes in the hidden layer. Nodes in a feedforward network are basically functions, that can be linear or nonlinear. The functions are adjusted by the network by changing the weight or bias of each node. Those changes in weight will lead to a change in the output value. There are no feedback connections in this network, the data just goes through the input layer to the output layer, layer by layer, and the outputs are fed back into itself. Feedforward neural networks can be used for forecasting data. To do this, the system must be trained by training samples, including training input and training output.

In order to choose an appropriate model for our research, we had to read and point out several variables and models that previous researchers were using in their articles. Those articles are suitable and contain more than 80 references.

The research of Tereza Sustrova (2016): “A suitable artificial intelligence model for inventory level optimization”, has its purpose of determine the right proposed method of neural network used to improve the supply chain [20]. The research mainly discusses using a neural network model to optimize the inventory level, improve the order system, and manage stock. For that purpose, the research uses Artificial Neural Network to forecast demand and determine quantity to order in order to optimize inventory level. Furthermore, the selection of input and output variables is significant to the model and this research of Tereza Sustrova uses 6 variables for input to determine 1 output variable, quantity to order. Those input variables are Current demand (IDact), demand in the next 3 months (ID3), demand in 3 months following after the 3-month order cycle (3-month delay) (ID33), current inventory level (II), purchase prices and transport costs (IT) [16]. The variables which we use in this study are mainly based on this research.

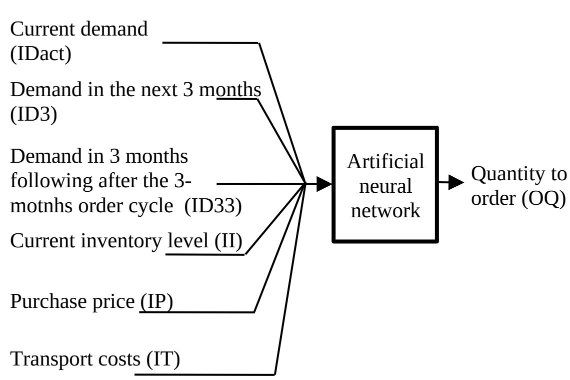

Figure 1: Diagram of the inputs and output of the constructed neural network (Tereza Sustrova, 2016).

The research of Zeynep Hilal Kilimci et el (2019): “An improved demand forecasting model using deep learning approach and proposed decision integration strategy for supply chain” mainly discusses which forecasting technique is sufficient for demand forecasting, including time series and regression methods, support vector regression model, feature reduction approach, deep learning methodology and a new final decision integration strategy [3]. In this research, we focus on the approach of using deep learning to forecast demand.

According to Zeynep Hilal Kilimci et el (2019), the deep learning approach is a machine learning technique that applies deep neural network architectures to solve various complex problems and the algorithm used in this study is a multilayer feedforward artificial neural network. In these algorithms, the system is trained with stochastic gradient descent using backpropagation [3]. The input variables used in this model that can be applied in our model are store number, stock in quantity, sales quantity, discount amount, customer count, max and min stock, average temperature, and weather condition.

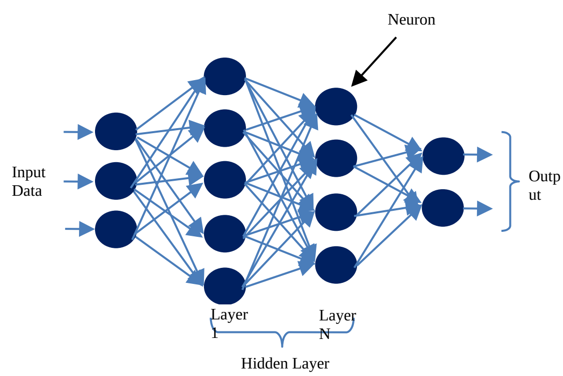

Figure 2: Diagram of the inputs and output of Deep Neural Network (Zeynep Hilal Kilimci et el., 2019).

The research of Luis Aburto (2007), named “Improved supply chain management based on hybrid demand forecasts”, discussed a hybrid intelligent system combining Autoregressive Integrated Moving Average (ARIMA) models and neural networks for demand forecasting. As the system helps to improve supply chain management in the retail industry, both of the methods used in this research show promising [21]. However, in terms of accuracy, neural networks outperformed ARIMA models and the proposed additive hybrid approach gave the best results. In our study, we use neural networks as a method to improve the supply chain management of the ice cream industry.
The research of Muniandy Sivaram (2016), named “Factors influencing consumption pattern of ice cream in Bengaluru market”, mainly discussed which factors affect ice cream consumption, especially expenditure on the consumption of ice cream [22]. Those variables are type of family, family size, annual family income, food habit, and monthly expenditure on non-food items. However, in our study, we did not apply those variables despite annual family income but the methods used in that research, including data collection, Logistic Regression Analysis,... are helpful for our study.
The research of Reza Toorajipour (2021), named “Artificial intelligence in supply chain management: A systematic literature review”, sought to identify the contributions of artificial intelligence (AI) to supply chain management (SCM) by reviewing several literatures.[23] In this research, many fields, including marketing, logistics, supply chain,... are discussed and AI approaches are applied to those fields. This research helps us to determine the use of AI in the supply chain and therefore find an appropriate model for our study.
In order to forecast demand in the ice-cream industry, variables need to be selected correctly and those variables need to have a major impact on ice-cream consumption. The research of Yan-Kwang Chen et al (2015): “A customer value analysis of Taiwan ice cream market: a means-end chain approach across consumption situations” described many factors which affect customer value in ice cream industries [24]. From this research, we recently pointed out 5 variables that may affect ice cream consumption. Due to the fact that the research focuses highly on ice cream brands, many variables in this research are inappropriate for us because we did not choose brands as a factor in our study.

1.4 Ice cream industry

There are many approaches to the definition of the ice cream industry. According to Wendell S. Arbuckle (2013), the ice cream industry has developed on the basis of an abundant economical supply of ingredients and is a high-volume, highly automated, modern, progressive, very competitive industry composed of large and small businesses manufacturing ice cream and related product [12]. The industry underwent a difficult period of adjusting to economic changes and to the establishment of product specifications and composition regulations. The industry struggled to adapt to economic shifts as well as the establishment of product requirements and composition regulations. Weather, fitness, taste, and impulse buying are four factors that affect ice cream demand, according to H. Douglas Goff’s (2013) theory [13]. The hot weather has a major impact on ice cream consumption. A preference for a balanced diet either discourages ice cream purchases or prefers low-fat or no-fat varieties. We primarily describe the ice cream industry in this study using H. Douglas Goff’s definition (2013) [13].

1.5 The impact of demand forecasting on the supply chain

A supply chain, according to Anirut Kantasa-ard (2020), is characterized as the entire process of producing and selling commercial products, including all stages from the procurement of raw materials to their distribution and sale [14]. Since it is influenced by so many unexpected factors, the supply chain is complicated and random. Many businesses are currently having trouble managing inventory and optimizing supply chains. A proper demand forecasting approach is a solution to this issue, but it is difficult to implement due to the varying demands of customers.
According to Zeynep Hilal Kilimci (2019) study, the organization has focused on the technique of analysis and prediction in order to reduce costs, increase efficiency, and optimize the supply chain model due to high competition in the market between retailers [3]. Overstock and out-of-stock are two big issues for retailers, as they have a negative impact on sales. The supply chain mechanism is all about cost, and retailers must manage the chain to minimize financial risk. The use of technology, especially artificial intelligence, is effective for demand forecasting and thus provides retailers with an appropriate tool.

1.6 Research hypothesis:

According to Thiago C.Silva (2019), the accuracy of a deep learning approach is calculated by comparing the predicted value with the real value or using the MAPE function. DL is an implementation of an artificial neural network, which mimics the natural human brain. However, in terms of data analysis, deep neural networks give a better performance than humans as they are capable of dealing with large amounts of data. As forecasting demands are basically forecasting data based on historical data, deep learning approaches show its potential. Therefore, we hypothesize the effect of DL accuracy on supply chain demand prediction.

Hypothesis 1*: The accuracy of Deep Learning approach affects positively on the outcome of demand forecasting in the supply chain.*

According to Thiago C. Silva (2019), demand forecasting is one of the main problems of the supply chain. It aims to optimize inventory, reduce costs, increase sales, profitability, and customer loyalty. For this purpose, historical data can be analyzed to improve demand forecasting by using different methods.
In this article, it can be seen that the accuracy of the model depends significantly on the historical data. So, the process of collecting data, the accuracy of the data given, and selecting the right variable for the data are remarkable. From there, we hypothesize the research on historical data influencing the results of demand forecasting of the supply chain.
In our research, we will test the network with 2 methods, with and without filtering variables to see if the selection of variables can affect the accuracy of the network.

Hypothesis 2: *The performance of the historical data collecting method and variable selection has a positive impact on the outcome of demand forecasting in the supply chain.*

The most common Willingness to pay (WTP) definition states: “WTP is the maximum amount an individual claims they are willing to pay for a good or service” (DFID, 1997). Customers may be unwilling to pay for a product, but they would do that instead of having nothing to use.
According to Wedgwood and Sansom (2003), there are two ways to approach and estimate WTP: revealed preferences and stated preferences. For the revealed preferences approach, we need to observe the price people pay for goods in different markets (e.g. buying water from vending machines or from neighbors, paying local taxes, etc); and personal spending on money, time, labor, etc to get goods. For a stated preferences approach, we need to ask people directly what they are willing to pay for future goods or services.
The issue of willingness to pay for products for policymakers and decision-makers is to ensure the implementation of that service is acceptable to the public. This may involve predicting what the user will be able and willing to pay for that product (Rohana Kamaruddin, 2012). Deep learning approaches can be used to forecast customer willingness-to-pay based on several factors. The accuracy of this model is potential because the model can analyze customer data with is relevant to their willingness-to-pay.

Hypothesis 3*: Deep Learning approach affects positively on the accuracy of customer’s willingness-to-pay forecasting.*

Based on the articles above, an approach of using a deep feedforward network is appropriate for demand forecasting and we might use these algorithms in our research to forecast. Our research target is the ice cream’s product of the grocery store named Vinmart. The inputs we are using in our model are shown in picture 3. The outputs we try to forecast in our model are quantity to order and customers’ willingness to pay. In order for the system to operate, a training sample is required and we used the historical inputs and outputs as training data, which is based on the business report, demographic data, and climate data of the store in the year 2018. The system is tested with the data of 2019 to see if the model is appropriate for demand forecasting. For this part, MatLab is a program we are using to operate the deep feedforward network and then we use the MAPE function to test if the model gives an accurate prediction.

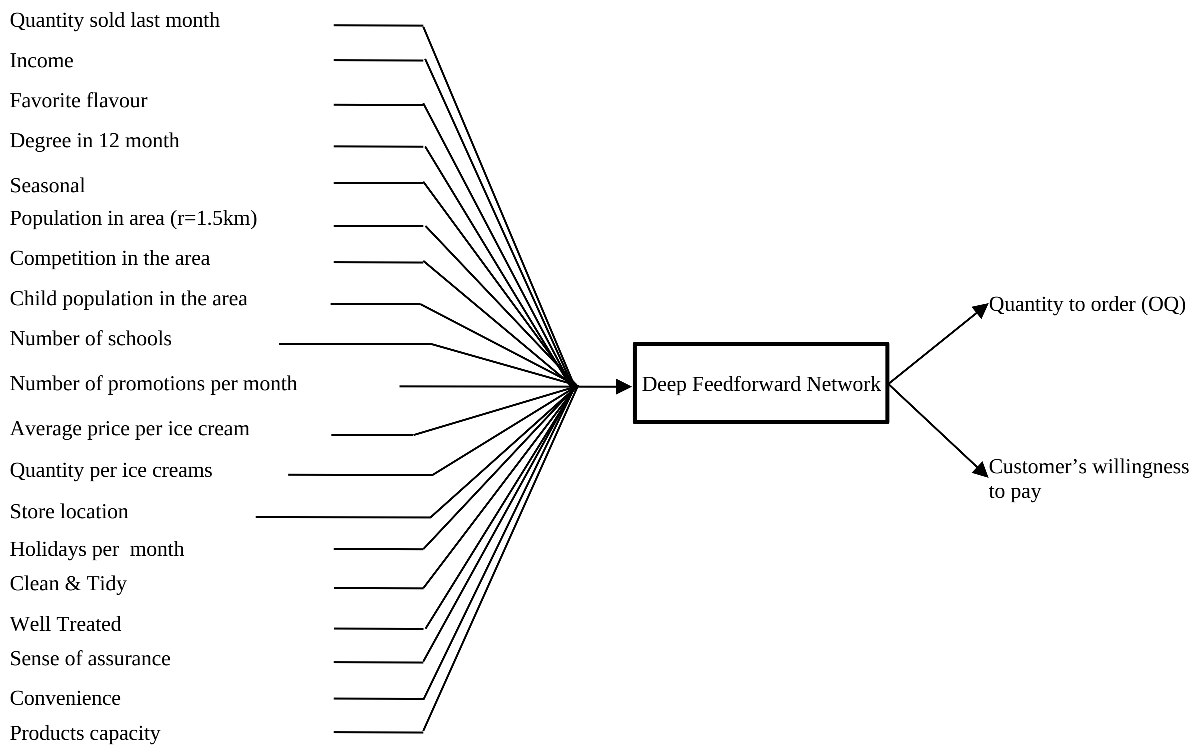

Figure 3: Diagram of the inputs and output of Deep FeedForward Network. Source: Zeynep Hilal Kilimci et al's study (2019).

4. METHODOLOGY

4.1. Data explain

There are 2 kinds of data that we focus on in this research which are training data and test data. For the training data, we use the data from the 2018 Vinmart store and the demographic information in the area in the same year. For a better result, more data should be collected and it is supposed to have data from as many past years as possible but since the stores we do research on were just opened in 2018, it is impossible to train the network with more data. However, it might be improving in the future as more data has been collected for a better network. For the test data, we use the data collected from 2019 of the same stores and the point of this data is to test the forecast value with the actual value so that we can see how accurate the network can reach.

Recently, our source data is related to weather data, demographic data, customers, and business reports of the store we do research on. The weather data is about temperature, and rain days per month which we found significant to the quantity of ice cream sold. For the demographic data, the data we do research on is population, number of children, and number of schools and we also do research about their feedback on Vinmart overall. Furthermore, a business report is significant to our model and it has to be accurate to test and train the network. Its impact on the quality or performance of the forecast system is the most significant in this model.

For our data collecting method, we collected data from 3 Vinmart grocery stores in Da Nang city and we separated them into 3 districts which are Hai Chau, Cam Le, and Ngu Hanh Son. These districts are different in demographics and those differences actually have an impact on the quantity of ice cream sold.

For Vinmart in Hai Chau, this grocery store is located in the center of Da Nang city and its population contributes 20% to the total number of the city. The average income of this district is also the highest. Moreover, there are a lot of competitors in this area that sell the same product with this store so it might have an impact on the revenue overall. However, the number of schools in this area is lower than Cam Le and Ngu Hanh Son districts. In Cam Le district, the population contributed only 8% to the total of the city while the average income in this area is low compared to Hai Chau district. However, this area has many schools and colleges so the number of students is high in this area. Furthermore, compared to Hai Chau district, there are only a few competitors in this area. Because of that, despite the low population, the quantity of ice cream sold in this district is potential due to these factors mentioned above. For Ngu Hanh Son district, the store we do research is located in a favorable location that has high traffic. Moreover, the number of schools and tourists in this area is remarkable. To sum up, there are some remarkable differences between the 3 stores so there will be a slightly difference in the result and throughout this test, we will find out if the network is working on several different factors between its elements.

4.2. Feedforward neural network

In this study, we use a feedforward neural network as an algorithm for deep learning. This network is considered useful in classification and forecasting. According to Tushar Gupta (2017), a feed forward neural network is an algorithm that contains several units, organized into 3 different layers: Input layer, hidden layer, and output layer. Data enters the network at the input layers and then is processed by nodes in a hidden layer. Nodes in a feedforward network are basically functions, that can be linear or nonlinear. The functions are adjusted by the network by changing the weight or bias of each node. Those changes in weight will lead to a change in the output value. There are no feedback connections in this network, the data just goes through the input layer to the output layer, layer by layer, and the outputs are fed back into itself. Feedforward neural networks can be used for forecasting data. To do this, the system must be trained with a training sample, including training input and training output.
The variable selection and the weight value are significant to the output of the network. In order for the network to perform well, the input variables have to be related to the output, and how much they affect the output is defined by weight. Unrelated input variables should be eliminated before running the network because it will cause inaccuracy and the weight value can be adjusted through testing.

4.3. Testing accuracy

The mean absolute percentage error (MAPE) is probably the most widely used forecasting accuracy [15]–[17]. MAPE is useful to estimate the accuracy of a forecasting method in order to determine how reliable a forecasting method is. However, several authors like Armstrong & Collopy (1992) and Makridakis (1986) have described MAPE as error distributions which means that MAPE can show how much error a forecasting method distributes overall [15], [18], [19].
For a given set of data, we will train and test the network on the forecasting task. We calculate the accuracy by the MAPE function we mentioned above. According to Lewis (1982), MAPE>50% means the forecasting method is inaccurate, MAPE<10% means the method is highly accurate, MAPE<20% means the method is good and if MAPE is from 20 to 50%, the forecasting method is reasonable at some points.
To sum up, we will use the MAPE function as a method to test the forecasting model's accuracy. The formula we use for MAPE is MAPE=mean*(abs*(At-Ft)/At). In which, At stands for actual value and Ft stands for Forecast value.
In this research, if MAPE<50%, the forecasting method will be reasonable and can be used at some point with concerns. Otherwise, we will describe the method as an inaccurate method that is not recommended to be applied. The lower the MAPE is, the more accurate the forecast model gets.

5. RESULTS AND DISCUSSION

5.1. Filtering variables

However, since not every variable has an impact on the output of the model, filtering variables is recommended. For this task, we use linear regression to estimate the significant level of the variables to the dependent one, which is quantity to order. The results of the linear regression are shown below:

Table 1. The regression results.

Variables	Beta	t	Sig.
Constant	-4040.513	-0.970	0.336
Temp	284.145	6.351	0.000
Holidays	-7.319	-0.117	0.907
Season	-570.069	-4.485	0.000
Child	-0.038	-0.147	0.883
School	-43.385	-0.587	0.559
Performance	417.388	4.648	0.000

As a result, the variables of population and number of promotions are not included since they don't bring additional significant information to the model. Due to F=21,184 and sig <0,05, we can conclude that the other variables such as store performance, season, temperature, number of holidays, number of schools, and children population can be analyzed by linear regression. As Adjusted R Square=0,63, we can conclude that the six variables mentioned above are responsible for 63% of the change in the dependent variable.
Based on the Coefficients table, we can see that the Coefficients's sig of temperature, season, and performance are lower than 0,05 so these variables do have an impact on the dependent variable. However, the same figure for the number of holidays, number of children and school are higher than 0,05 which means that they don't affect the variable of quantity to order.

To sum up, it seems to be that only temperature, season, and store performance have a significant impact on the quantity of ice cream sold in our model. However, to see if filtering variables are important for a deep feedforward neural network on forecasting tasks, we will test the network in two ways: without filtering variables and with filtering variables. It means that although the number of holidays, number of children, and school don't actually affect the output, we still apply those variables on the network to test how the network performs without filtering variables.
Furthermore, we also test the network to see if it is able to predict how much customers are willing to pay for an ice cream based on their personality and how they feel about the stores' service. The purpose of this network is to help managers select which price of ice cream to attract more customers. However, the variables and dataset will be completely different from the network that forecasts the quantity of ice cream sold.
To perform this task, we gave a questionnaire to the local people and asked them about how they feel about Vinmart's service and how they are willing to pay for an ice cream there. We receive up to 60 answers and in this research, we will use 40 answers to train the network and 20 answers to test the accuracy.

4.2 Experiment Results

Our forecasting model uses the data we collect as we mentioned on the setup. We will compare the forecasting result with the actual one to estimate the accuracy of the model. However, there are several results as we use different estimated approaches.

a. **The integration method:** Without filtering variables:

For this approach, we apply the feedforward network to forecast demand for 3 different stores at the same time. Moreover, for more details, we test the network by adjusting the weight and we find out that the accuracy, calculated by the MAPE function, is different between the same network but with different weights. For the first network with 4 weights, the result we receive is not promising since MAPE= 56,71% which means that the forecasting method is inaccurate with 4 weights. However, while we adjusted the weight to 2, there was a huge change in the accuracy of the method since MAPE was reduced and reached 42,257%. For this number, the forecasting method is reasonable but it is not considered a good forecasting method.
With filtering variables, after we eliminate the variables that we prove are not significant to the model, the forecasting methods become even more inaccurate with MAPE with w=4 is 67,39% and MAPE with w=2 is 57%. However, as the variables are filtered, the variables that still remain significant like the weather condition, season, and store performance will have more impact and under these circumstances, the weight of the network has to be adjusted to a higher level since the input becomes more significant. After testing with several weight values, the network with weight=12 is the most appropriate with MAPE=40,232%. In conclusion, it is a reasonable forecasting method and we can conclude that filtering variables can increase the accuracy of the model but the weight of the network has to be adjusted since the remaining variables become more or less significant.

b. Forecasting on each store:

For this approach, we will train the network with the data from 3 stores, which will be separated from each other so that there will be 3 networks corresponding to each store. We will use the same approach as the integration method as we will adjust the weight of the network to see the difference.

b.1. Applying the forecasting method on the store Vinmart in Cam Le:

Without filtering variables:

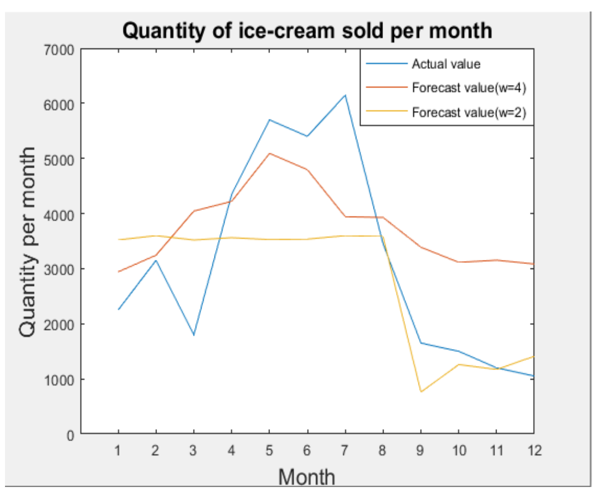

Figure 4: Quantity of ice-sream forecast in Cam Le

As we use the forecast method on the store in Cam Le, the results we receive are shown in the table above. The table shows that there is a big difference between the two methods as it gives different results and it is recommended to compare the methods in order to find the optimal way to forecast the quantity of ice cream in this store. However, it is difficult to compare the performance of 2 different forecasting methods through the table so the MAPE function is recommended. For the method with weight = 4, the MAPE we receive is 66,88% so the method is not as accurate as it should be. For the method with weight = 2, the MAPE we receive is 34,09% so the method is reasonable and the forecasting is accurate at some point in the data. To sum up, in this store, by using a lower weight, the performance of the forecasting method seems to improve, which means that the quantity of ice cream sold in this store is not highly dependent on its variables. (Weights in the feedforward network show how much the variables matter in the model). A Feedforward network with weight=2 is recommended to forecast the quantity of ice cream in this store because it gives a better result overall.

With filtering variables:

After we eliminate variables that don't have a significant role in the model and test the network with several weight values. For the network with weight=4, the MAPE we receive is 28,5%, which means that the forecasting method in this case is quite accurate. In fact, compared with the network with the same weight but without filtering variables, the network shows a potential increase in accuracy. However, when we tried to adjust the weight to 12 as the same as the integration method we used above, it seemed to be inappropriate on this method since MAPE reached 117%.

Because of that, we can not apply the same network we use for the integration method to another method or to a different target.

Because of that, we conclude that the weight of the network has to be adjusted for each source of data for the best performance. The reason for that is the input has a different meaning to the dependent variables between each source of data.

b.2. Applying the forecasting method on the store Vinmart in Hai Chau:

Without filtering variables:

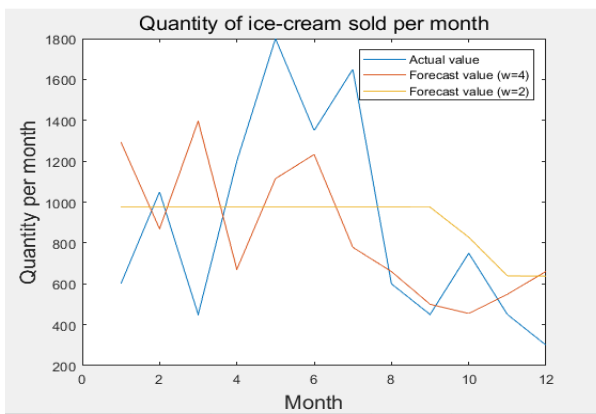

Figure 5: Quantity of ice-cream forecast in Hai Chau

When we apply the forecasting method to the store in Hai Chau, the results we receive are different between the same neural network with weight=4 and weight=2. However, the result from the network with w=2 stays constant from January to September while the actual value fluctuates wildly in this period, this is not a good sign which proves that the method is inappropriate. As for the results, on this store, the results we receive are deficient since the MAPE is over 50% on both methods. It is not recommended to use the method on this store at the moment based on the low performance it gives.

For better results, more data has to be collected and the selection of variables has to be more appropriate to forecast demand by the feedforward neural network in this store since we find many errors occurring at the moment. Unfortunately, when we try to overcome the problem by collecting more data from this store in the years 2017, and 2016, it is impossible because the store was just opened in 2018.

With filtering variables:

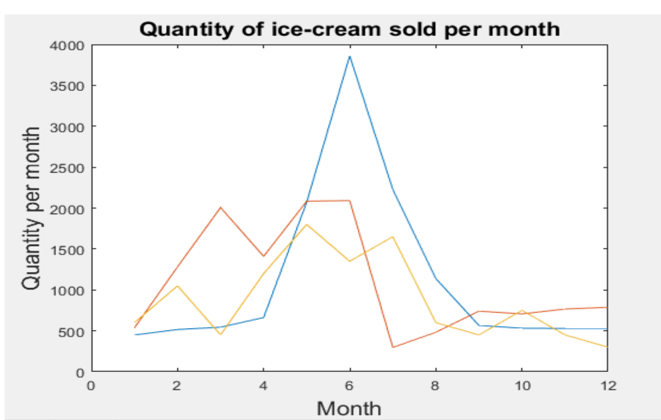

Figure 6: Quantity of ice-cream forecast in Hai Chau

At this store, after we eliminate the unrelated variables, the MAPE value seems to decrease on both networks with w=4 and w=2. However, the accuracy of this method on the Hai Chau Vinmart store is still not considered trustworthy since MAPE values are always > 50% on every network.
We can conclude that our forecasting method by using a feedforward network has not worked on this store recently.
To sum up, both forecasting methods with and without filtering variables are not working on this store.

b.3. Applying the forecasting method on the store Vinmart in Ngu Hanh Son:

Without filtering variables:

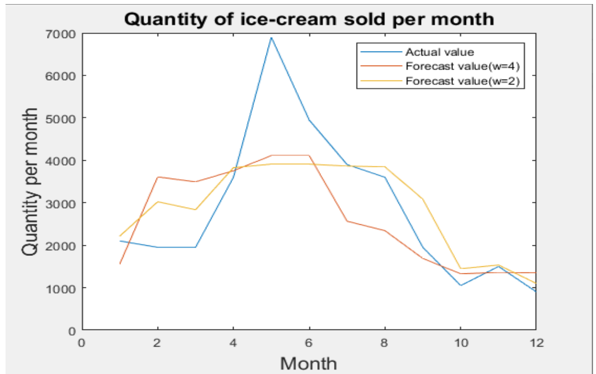

Figure 7: Quantity of ice-cream forecast in Ngu Hanh Son

These forecasting methods fail to predict the huge increase in the number of ice cream sold from April to June, 2019. However, it is said that the unexpected increase in this period is too random and can not be explained by any variables we found.
Overall, the method seems to be acceptable in this store while the MAPE while weight=4 is 34,96%, and the MAPE while weight=2 is even better with 25,397%. Because of this, the forecasting method is trustable in this store.

With filtering variables:

After eliminating unrelated variables, the network performs poorly with the weight=4 and 2 since MAPE is over 50%. However, since we adjust the weight to 12, the network performs way better with MAPE= 28,06% which shows that the forecasting method is reasonable.
However, in conclusion, there is no significant difference between the methods with and without filtering variables but forecasting the quantity of ice cream by using a deep feedforward neural network can be used on this store.
In conclusion, a deep feedforward network can be applied to forecasting tasks but filtering variables is recommended and the weight of the network needs to be adjusted based on how the input data affects the output. To take this further, we will do an experiment on how a feedforward neural network can be used to forecast willingness to pay.
Using Deep Feedforward Neural Network to forecast customers' willingness to pay. we point out many variables considered to affect customers' willingness-to-pay. From the variables we selected, we collect data from questionnaires and forecast the results with a feedforward neural network (w=8). We trained the network with the data we received and as a result, the MAPE we received is 38,9% which means it is a reasonable forecasting method. Below is the model for the neural network that combined 8 hidden layers,9 inputs, and 1 output.
As we forecast customers' willingness-to-pay, we test the system with both the training and the test data. We received 44 answers through questionnaires so we are trying to forecast how much customers are willing to pay for an ice cream from 44 answerers. The chart below shows the difference between the forecasting value and the real value.

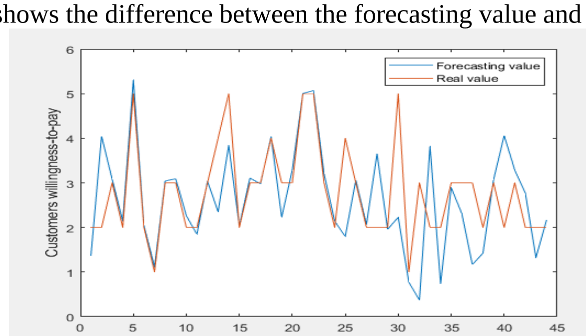

Figure 8: Quantity of ice-cream forecast in Ngu Hanh Son

6. CONCLUSION

7. In the retail industry, demand forecasting is one of the most important tasks in the supply chain to optimize stock, reduce costs, increase sales, profitability, and customer loyalty. To perform this task and overcome the issue, the

method of using Deep Learning is approached to analyze and learn complex interactions and patterns from historical data to make predictions about order quantity and customers' willingness to pay. Based on the theoretical bases that we have presented, in this study, the Deep Feedforward Network model is proposed to test and implement actual data obtained from VinMart and other retail stores. Therefore, Chapter 1, the theoretical basis, has made the foundation for our analysis and research in the future.

Based on the experiment result, the methods we use have their own contributions and limitations. As we use the neural network to forecast demand in 3 different stores at the same time on the same network, the result we receive is promising but we would not recommend it. The reason is that using the network on each store, one network is trained with only the source of data from only one store and location, might give a better result overall. Filtering variables can help increase the accuracy of the data but it is quite dependent on the target. Moreover, using the same network for every store is appropriate but the weight values have to be adjusted due to the fact that the input's significant level is different between those stores. In this research, the accuracy of a feedforward neural network on forecasting tasks seems to be reasonable but not a good method overall. However, there is a struggle that we have to deal with is that the stores can only provide data down to 2018 since it was opened that year while it is supposed to collect more past year data to train the network in order to get a good performance.

In this research, we discussed some recent applications of Deep Learning algorithms for the Demand Forecasting model. The algorithm we mainly used in this research is Deep Feedforward Network, generated by MatLab, which can forecast the future output based on training data and recent (test) data. The output its given is logical and promising since we test the model by comparing the test output with the forecasting output in a stimulating environment and achieving low errors. However, the accuracy of this forecasting model might be deficient due to many factors, considering the lack of data collected, inaccurate data, irrelevant input, and unexpected instances. Those factors significantly affect the model, though, it is hard to be pointed out. In the context of this problem, it is significant to collect the correct data and identify the appropriate input in order to achieve the best performance in the model. For unexpected instances, which can cause a major impact on the forecasting method, the model is considered unreliable in this circumstance since this network is not able to learn and adapt to an occasion which is never happened in the past unless we find the correlation coefficient between the unexpected instance's input and the possible output. If only the problems are solved, the model can predict the right quantity to order and therefore, optimize the supply chain.

Also in this research, we use a deep learning approach to forecast the demand for ice cream, including quantity to order and customer's willingness to pay. The inputs of this model are shown in the table above. Except for some inputs that can be calculated and presented in numbers, inputs that are based on customers' value according to Chen are collected by quantitative research and presented on a scale from 1 to 7. Those inputs will be conducted to predict the customer's willingness to pay and we will test the model by comparing the predicted value with the actual customer's willingness to pay, which is collected in the questionnaires. If the predicted value is accurate, we will use the data collected and predicted to find the connection between customers's value and how much they are willing to pay.

The accuracy of this model is highly dependent on data so it is only recommended for companies with a superior database and effective data management. The cost for the infrastructure to generate the model is affordable even when the use of it on demand forecasting is huge as long as the data are appropriate. This model shows directions to solve the main problem in the operation of the supply chain, especially overstock and out-of-stock.

Furthermore, big data analysis tasks are important and require the ability to work with high amounts of data or data patterns. Those big data analysis tasks can be used on demand forecasting and recently can be enhanced by the DeepLearning approach according to Maryam et al [25]. As we define the connection between customers's value and how much they are willing to pay through the model by using this model, it is appropriate for businesses to identify with customers's value to focus on.

The limitations of this study are as follows:(1) We only selected Vinmart stores in DaNang city as our subjects. Future researchers are advised to conduct research across many stores and places with different climate conditions since ice cream consumption is highly dependent on climate; (2) We did not classify ice cream by brand so brand factors are not included in the research. Future researchers might use brand factors as variables and therefore, improve the accuracy of the model; (3) We only use a deep feedforward network as an approach to deep learning. Thus many Deep Learning approaches can be used for demand forecasting, but a better model for this purpose can be missing. Future researchers can use the LSTM network, recurrent network, or any different algorithms to identify which one gives the most accurate output; (4) For our network on forecasting customers' willingness-to-pay, the data we test the network with is not wide enough. More data has to be collected and target wider.

8.

9. REFERENCES

1. P. Hà, "Phát triển trí tuệ nhân tạo (AI) tại Việt Nam: Thực trạng, kinh nghiệm quốc tế và xu hướng phát triển | Học viện Cảnh sát nhân dân," Nov. 14, 2019. http://hvcsnd.edu.vn/nghien-cuu-trao-doi/dai-hoc-40/phat-trien-tri-tue-nhan-tao-ai-tai-viet-nam-thuc-trang-kinh-nghiem-quoc-te-va-xu-huong-phat-trien-5675 (accessed May 29, 2021).
2. McKinsey&Company, "Smartening up with Artificial Intelligence (AI) - What's in it for Germany and its Industrial Sector?," Digit. McKinsey, p. 52, 2017, [Online]. Available: https://www.mckinsey.com/~/media/McKinsey/Industries/Semiconductors/ Our Insights/Smartening up with artificial intelligence/Smartening-up-with-artificial-intelligence.ashx.

3. Z. H. Kilimci et al., “An improved demand forecasting model using deep learning approach and proposed decision integration strategy for supply chain,” Complexity, vol. 2019, 2019, doi: 10.1155/2019/9067367.
4. K. G. Kim, “Deep learning book review,” Nature, vol. 29, no. 7553, pp. 1–73, 2019.
5. A. Oppermann, “What is Deep Learning and How does it work? - Towards Data Science,” Medium, Nov. 2020, Accessed: May 29, 2021. [Online]. Available: https://towardsdatascience.com/what-is-deep-learning-and-how-does-it-work-2ce44bb692ac.
6. T. Gupta, “Hardware Implementation of Neural Network for Image Classification.” State University of New York at Buffalo, 2020.
7. M. Webster, “Effective demand for rural water supply in South Africa,” procedings WEDC Conf., pp. 25: 426-429, 1999.
8. W. P. Van Amstel, Trends in supply chain management, vol. 2018-June, no. 7–8. 2018.
9. J. White, “Evaluation synthesis of rural water & sanitation projects.” pp. iv, 74 p., May 01, 1997, Accessed: May 29, 2021. [Online]. Available: https://www.ircwash.org/resources/evaluation-synthesis-rural-water-sanitation-projects.
10. D. W. Pearce, Macmillan dictionary of modern economics. Macmillan International Higher Education, 1986.
11. G. Zhang, B. Eddy Patuwo, and M. Y. Hu, “Forecasting with artificial neural networks: The state of the art,” Int. J. Forecast., vol. 14, no. 1, pp. 35–62, 1998, doi: 10.1016/S0169-2070(97)00044-7.
12. W. S. Arbuckle, Ice cream. Springer, 2013.
13. H. D. Goff and R. W. Hartel, Ice Cream - H Douglas Goff, Richard W Hartel. 2013.
14. A. Kantasa-ard, M. Nouiri, A. Bekrar, A. Ait el cadi, and Y. Sallez, “Machine learning for demand forecasting in the physical internet: a case study of agricultural products in Thailand,” Int. J. Prod. Res., pp. 1–25, Nov. 2020, doi: 10.1080/00207543.2020.1844332.
15. J. S. Armstrong and F. Collopy, “Error measures for generalizing about forecasting methods: Empirical comparisons,” Int. J. Forecast., vol. 8, no. 1, pp. 69–80, 1992, doi: 10.1016/0169-2070(92)90008-W.
16. P. Goodwin and R. Lawton, “On the asymmetry of the symmetric MAPE,” Int. J. Forecast., vol. 15, no. 4, pp. 405–408, 1999, doi: 10.1016/S0169-2070(99)00007-2.
17. L. Ren and Y. Glasure, “Applicability of the revised mean absolute percentage errors (MAPE) approach to some popular normal and non-normal independent time series,” Int. Adv. Econ. Res., vol. 15, no. 4, pp. 409–420, 2009, doi: 10.1007/s11294-009-9233-8.
18. C. Beaumont, S. Makridakis, S. C. Wheelwright, and V. E. McGee, “Forecasting: Methods and Applications,” J. Oper. Res. Soc., vol. 35, no. 1, p. 79, 1984, doi: 10.2307/2581936.
19. F. Collopy and J. S. Armstrong, “Another Error Measure for Selection of the Best Forecasting Method: The Unbiased Absolute Percentage Error,” Int. J. Forecast., vol. 8, no. 2, pp. 69–80, 2000, [Online]. Available: http://hops.wharton.upenn.edu/forecast/paperpdf/armstrong-unbiasedAPE.pdf.
20. T. Šustrová, “TRENDY EKONOMIKY A MANAGEMENTU TRENDS ECONOMICS AND MANAGEMENT A Suitable Artificial Intelligence Model for Inventory Level Optimization Tereza Šustrová: A Suitable Artificial Intelligence Model for Inventory Level Optimization,” vol. 25, no. 1, pp. 48–55, 2016, [Online]. Available: http://dx.doi.org/10.13164/trends.2016.25.48.
21. L. Aburto and R. Weber, “Improved supply chain management based on hybrid demand forecasts,” Appl. Soft Comput. J., vol. 7, no. 1, pp. 136–144, 2007, doi: 10.1016/j.asoc.2005.06.001.
22. Sunilkumar, M. Sivaram, and P. K. Dixit, “Factors influencing consumption pattern of ice cream in Bengaluru market,” Indian J. Dairy Sci., vol. 69, no. 4, pp. 492–497, 2016.
23. R. Toorajipour, V. Sohrabpour, A. Nazarpour, P. Oghazi, and M. Fischl, “Artificial intelligence in supply chain management: A systematic literature review,” J. Bus. Res., vol. 122, no. September 2020, pp. 502–517, 2021, doi: 10.1016/j.jbusres.2020.09.009.
24. Y. K. Chen, P. S. Tsai, and F. R. Chiu, “A customer value analysis of Taiwan ice cream market: a means-end chain approach across consumption situations,” Springerplus, vol. 4, no. 1, pp. 1–17, 2015, doi: 10.1186/s40064-015-1564-8.
25. M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, “Deep learning applications and challenges in big data analytics,” J. Big Data, vol. 2, no. 1, pp. 1–21, 2015, doi: 10.1186/s40537-014-0007-7.