

Optimized Warehouse Management of Perishable Goods for a Food Delivery Company

Snehalatha S¹, Suhirtha R²

1. Student, R.M.D Engineering College
2. Student, R.M.D Engineering College

Abstract

In recent years, food wastage becomes the major problem of the world and researchers indicate that 20-60% of the total production is lost in the food supply chain. Throughout the supply chain, fresh food companies face the challenge of handling both product quality due to the perishable nature of the product and its cost. A Machine Learning system and Cloud both integrated together taking into account supply chain constraints such as supplier delivery times and minimum or maximum order quantities. Based on all these considerations, it then generates order proposals for the entire product range every week. This is done with the use of Random Forest Regressor method. The aim of this project is to present a predictable order of the perishable supply chain management modeling and optimization approach focusing on loss minimization along the supply chain. We focus particularly on perishable goods and this project is analyzed based on the accuracy of the train and test data.

Keywords: Machine Learning, Cloud, Random Forest Regressor, Python

1. Introduction

The need to feed an ever increasing world population makes it necessary to reduce the tons of perishable waste along the food supply chain. About one third of global fresh fruits and vegetables (FFVs) are thrown away due to their quality. The particular aspect that must be considered in the competitive management of fresh food supply chains is concern for the deterioration in product quality. Maintaining high food quality standards is important for product market value, customer happiness and in turn for the long term reputation of the organization.

It has been a formidable task to handle the perishable food supply chains due to its short lifespan and the possibility of spoilage of the product due to its deterioration nature. All these components can cause a substantial amount of shortage of food items and retail loss. With the use of Machine Learning and its algorithms we have created a way for the optimization of the goods so that the wastage of perishable goods can be minimized to a maximum level.



Fig 1. Current Scenario

The solution to this problem we have proposed concentrates on both Cloud for collecting dataset and Machine Learning through Python to predict the amount of goods to be ordered in a range of one week. This is possible by collecting and storing the data of each order a customer makes for a range of one year in Cloud and then processing this data using Machine Learning Random Forest Regressor algorithm to predict the quantity of goods to be ordered in the next week and also for the consecutive ten weeks.

With the outreach of online shopping on the advancement in the delivering services many customers started to buy perishable products on Business to Customer platforms. Thus we have implemented and integrated this idea of webpage for online shopping and machine learning and cloud-based services.

By doing so, we will be able to predict the demand of products for the following week, buy and store only the required amount of perishable goods in the warehouses so that the wastage of food and money can be avoided at the same time. Thus, this helps both the retailer and customer in a way that the loss for retailers will be minimized with customer satisfaction as the customers get good quality food products.

2. Literature Survey

One of the approaches for a good supply chain management outcome is first-expired-first – out (FEFO), which was first introduced at the end of the 1980s. The basic concept is to implement stock rotation in such a way that the leftover storage time of each item is best matched to the remaining transport duration options, to reduce product waste duration transportation and provide product continuity at the store.

The management of fresh food supply chain, including deep frozen food has received increasing research attention in recent years. Ahumada and Villalobos (2009) presented an

overall study of agro-food supply chain. Labuza (1982) reviewed the significance of preserving and maintaining the freshness of food products was reviewed in industry and academic sectors. Rong et al. (2011) introduced an approach that integrates food quality in decision making for production and distribution in a food supply chain. They also determined product deterioration by time and temperature. Few scholars took customer satisfaction into consideration. If the customer was not convinced, a penalty cost would be generated for the spoilt food. Amorim and Almada-Lobo investigated correlation between distribution scenarios and the cost freshness trade-off.

3. Theoretical Analysis

The growth in data volume and diversity has led to data sets larger than ever before. To manage and evaluate these new and potential valuable data sets, new methods and application have been developed in the form of predictive analytics. One of these methods is Machine Learning (ML). ML has helped us to predict the amount of goods the retailer needs to store in his warehouse so that his perishable goods do not deteriorate. This was possible as a prediction is ideal for mobile applications, websites and other applications that use results interactively.

3.1 Block Diagram

The below block diagram shows the schematic representation of the process that takes place.

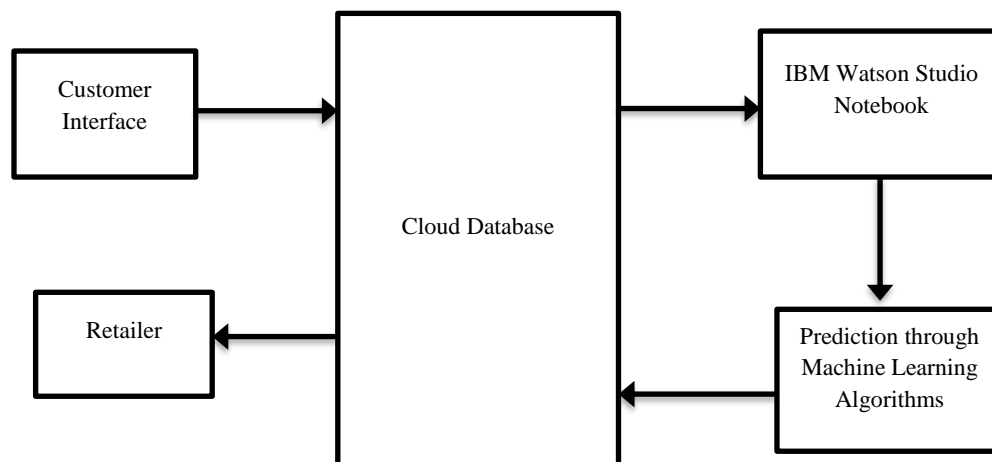


Fig 2. Block Diagram for Optimization of Warehouse

The above shows the flow of work and modeling of the process. From the user the data is collected and is stored and processed and then sent to IBM Watson Studio Notebook Platform where the prediction is done for the following week and next ten weeks from the previous week's data and the prediction is again stored in cloud which later on the retailer can access and view it.

3.2 Software Designing

The data set required to predict the amount of goods will be purchased during the next full week by machine learning is done by getting the data from the customers. This is made possible by creating a website for the retailer where he will showcase all the goods available in his warehouse. Each time a customer makes an order, the data of goods and its quantity is sent to the cloud where they are stored. After a time span of one week the data set collected is sent to the IBM Watson Studio Notebook, where machine learning is applied and the amount of goods to be ordered for the following week is predicted. The retailer can easily understand and makes a profit by ordering a correct amount and selling them without any deterioration.

4. Experimental Investigations

The data was collected from Kaggle platform as the work hasn't come into play. The entire dataset was used to predict using many algorithms. The dataset consists of 913000 rows and 7 columns namely store number, item number, year, month, day of week, week of year, and day of week. There were totally 50 items and 10 stores.

The table 1 shows the prediction accuracy of the entire dataset (in percentage) using various algorithm and their remarks was noted. It is clear from the table that SVR regression takes a lot of time to produce an output since the dataset has nearly 913000 rows.

Algorithms	Training Accuracy	Testing Accuracy	Remarks
Linear Regression	9.386	9.314	Has a very low accuracy
Polynomial Regression (degree 2)	18.43	18.5	Has a very low accuracy
Polynomial Regression (degree 3)	18.67	18.76	Has a very low accuracy
Support Vector Regression (kernel = rbf, C= 9)	There was no output for a runtime of 1-day	There was no output for a runtime of 1-day	Takes too long to produce output
KNN (C = 7)	51.12	34	Comparatively high accuracy and takes less time
Decision Tree Regression (Criterion = mae)	27.98	27.92	Has a nearly 50% accuracy only and also takes time
Random Forest Regression	46.3	45.76	Has a higher accuracy than other models

Table 1: Accuracy for the entire dataset

As the database was huge with nearly 913000 rows, which will be similar to the real-time data, it did not perform well both in time competency and as well as in accuracy determination. Thus the database was filtered and predicted with each item number and also with each store number.

The below table, table 2, shows the prediction accuracy of the dataset filtered shop wise (in percentage) using various algorithm in Machine Learning. It is clear from the table that all the algorithms gives very low accuracy in the range of 10 to 50% which will not be suitable for a large-scale company.

Store No. / Algorithm	Linear Regression		Polynomial Regression (degree 2)		Polynomial Regression (degree 3)		Support Vector Regression (kernel = rbf, C=9)		KNN (C = 7)		Decision Tree (Criterion = mae)		Random Forest Regressor (Criterion = mae)	
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
1	11.37	10.58	21.94	21.58	22.19	21.9	32.67	31.24	43.11	23.79	32.25	31.42	48.24	46.95
2	11.49	10.86	22.33	22.22	22.59	22.57	33.09	31.83	43.87	24.73	31.9	31.28	49.29	48.39
3	11.45	10.71	22.19	21.89	22.46	22.23	32.93	31.72	43.52	24.47	33.19	32.66	48.67	48.05
4	11.39	10.63	22.01	21.84	22.27	22.2	32.7	31.58	43.49	24.15	32.9	32.09	48.36	47.55
5	10.96	10.34	21.35	21.36	21.6	21.64	31.8	30.78	42.46	23.25	32.19	31.84	47.6	46.81
6	10.94	10.26	21.36	21.28	21.63	21.59	31.93	30.77	42.85	23.47	31.42	31.07	47.59	46.85
7	10.91	10.23	21.27	20.92	21.52	21.2	31	30.75	42.32	23.03	31.64	31.15	46.8	46.01
8	11.49	10.74	22.19	22.09	22.46	22.43	33.05	31.82	43.77	24.51	33.44	32.51	48.83	47.75
9	11.39	10.84	22.07	22.03	22.32	22.38	32.72	31.73	43.33	24.5	33.02	32.31	48.83	47.92
10	11.3	10.83	22.1	22.05	22.37	22.38	32.93	31.67	43.67	24.23	33.43	32.68	48.41	47.79

Table 2: Accuracy for each Store number

So prediction of Item wise was also carried on which is given in the table 3. This shows the predicted accuracy (in percentage) for various algorithms tested for first ten item number alone. It is clear from the table that SVR, KNN and Random Forest Regression algorithms give an average accuracy of 60 to 80%.

Item No. / Algorithm	Linear Regression		Polynomial Regression (degree 2)		Polynomial Regression (degree 3)		Support Vector Regression (kernel = rbf, C=9)		KNN (C = 7)		Decision Tree (Criterion = mae)		Random Forest Regression (Criterion = mae)	
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
1	19.16	16.66	38.17	35.91	38.65	36.61	62.42	60.22	67.87	55.71	48.92	48.21	63.47	60.8
2	23.33	23.13	46.51	45.95	46.87	46.5	76.39	75.29	80.13	72.1	59.14	58.84	77.04	75.49
3	21.94	20.18	42.73	42.26	43.24	42.72	70.39	68.67	74.67	64.7	54.46	53.16	70.98	68.51
4	18.09	18.26	37.79	37.14	38.33	37.64	62.11	59.91	67.67	54.81	48.63	48.13	63.85	61.18
5	18.43	17.35	36.64	33.94	36.93	34.28	58.96	55.93	65.18	50.46	45.23	43.3	60.96	57.33
6	23.58	22.5	47.02	46.27	47.63	46.56	76.8	74.99	80.6	72.01	59.18	59.09	76.52	74.38

7	23.31	21.91	46.84	46	47.47	46.52	76.3	75.14	80.28	71.63	58.48	59.04	76.16	75.82
8	24.55	23.5	48.34	47.68	49.01	48.25	79.05	78.02	82.59	74.77	61.16	61.38	78.80	78.25
9	23.01	21.74	45.53	45.63	46.06	46.15	74.26	73.97	78.35	70.25	58.26	57.57	73.71	72.18
10	24.61	23.16	48.16	47.39	48.82	47.8	78.82	77.35	82.18	74	60.74	60.81	79.01	78.04

Table 3: Accuracy for each Item Number

From the above two tables, table 2 and table 3, depicts that the accuracy was poor while predicting the number of goods needed in each store. So the item number was considered as the key filter for the prediction of the number of goods needed for the next or the following week. Also in the table 3, SVR, KNN and Random Forest Regression algorithms has the highest accuracy when compared to other algorithms. Among KNN and SVR, SVR consumes more time than KNN. But Random Forest Regression is more accurate and satisfying algorithm for this problem. So this model was opted for Random Forest algorithm.

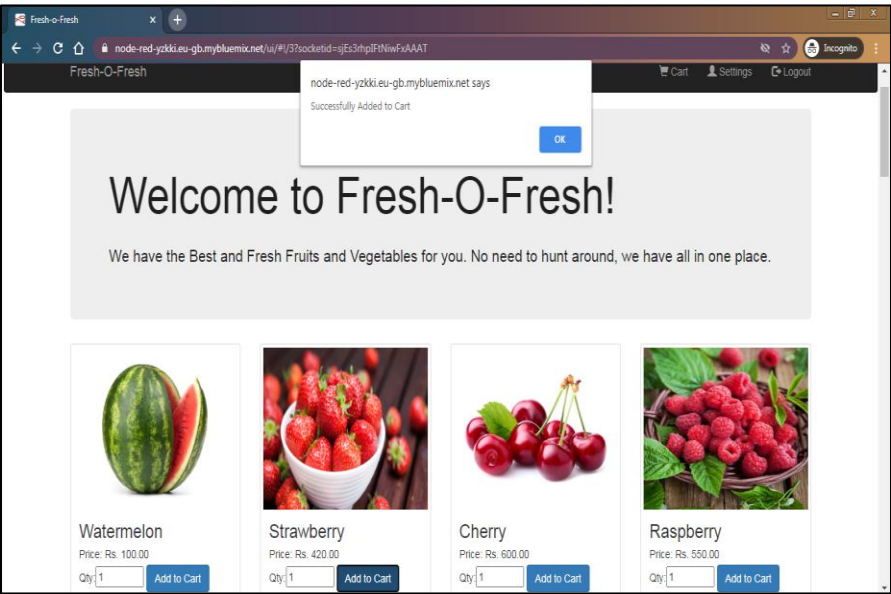


Fig. 3, User Interface

Fig 3 depicts the design of the products page designed for the customers which will get the buying information and store in Cloudant as a dataset.

Fig. 4 shows the input the retailers give (on the left side), the prediction of the week (on the top) and the prediction for next ten weeks (on the right). In this way the retailer will be informed on the trend of the upcoming weeks.

Important info:	
Prediction for 1st week	1670.8
Prediction for 2nd week	1670.8
Prediction for 3rd week	1670.8
Prediction for 4th week	1666.6
Prediction for 5th week	1666.6
Prediction for 6th week	1666.6
Prediction for 7th week	1666.6
Prediction for 8th week	1666.6
Prediction for 9th week	1666.6
Prediction for 10th week	1448.5

Fig 4 Prediction form for retailer

5. Steps and Flowchart

Step 1: Start

Step 2: Collecting the Login Information from Customers and storing them in Cloudant

Step 3: If a user logs in, checking if the email id and password in cloud redirect to Products page

Step 4: Otherwise, print Incorrect Password

Step 5: Collecting the Goods Information from the users and storing them in Cloudant and updating it each week with fresh data of the week and storing it in another database

Step 6: This information is filtered and creating data visualization

Step 7: Applying Machine Learning Algorithm

Step 8: Report on prediction sent to the retailer's UI

Step 9: Difference between the values in Step 5 & 6 for calculating loss

Step 10: Retraining

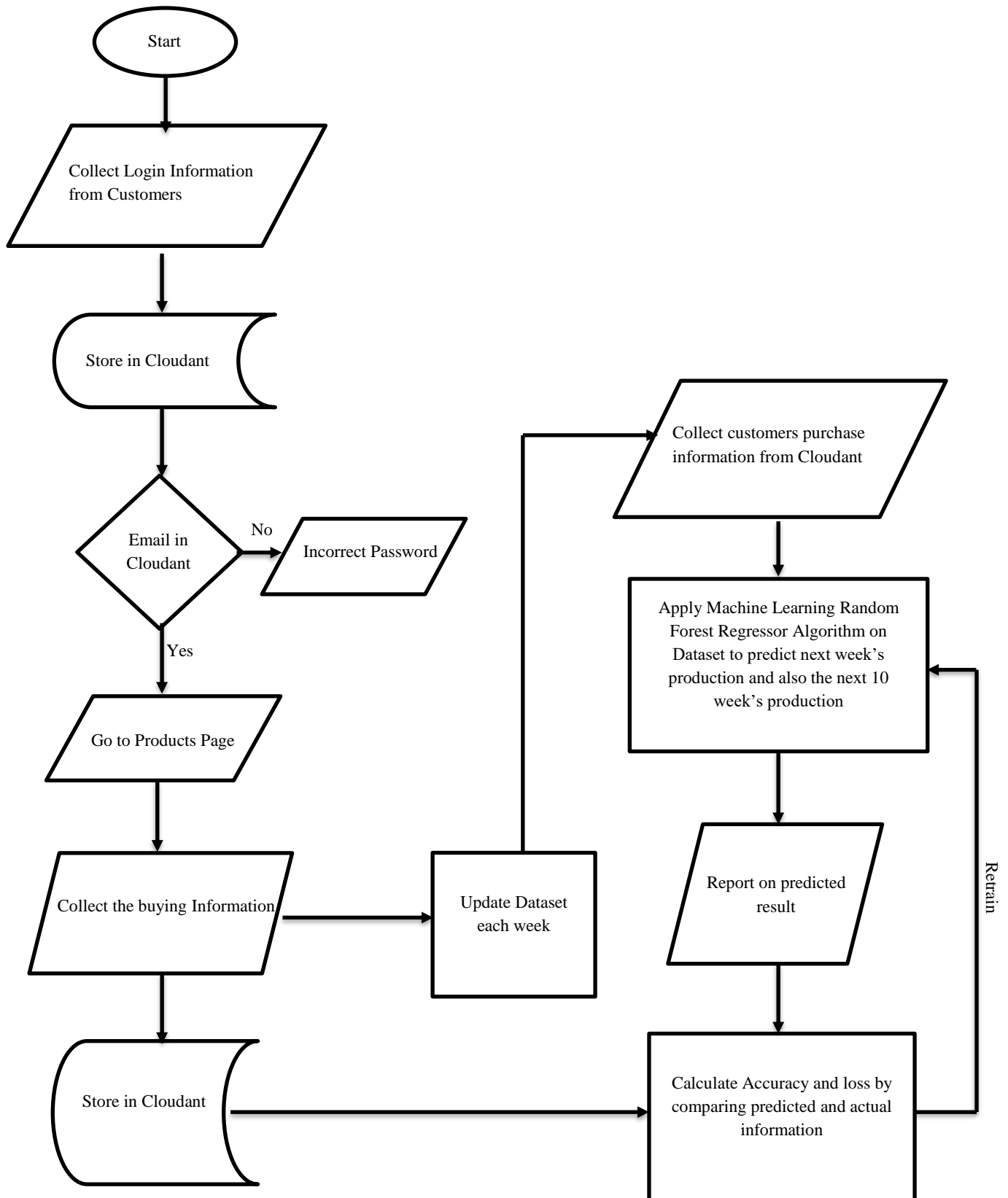


Fig 5 Flow chart

6. Result

The model is developed in standard conditions for testing purposes. The website works in the same way a normal online shopping website works, but with an extra advantage of storing all the user information and also their buying information in the cloud. The payment page is developed with highly secured.

The model is deployed in such a way that when the retailer requests the prediction for the start of the week, he gets the predicted value for that particular item for that week and also for the next ten weeks. So that with this, he will be able to identify the trend of the item in his shop and the way he can improve the sales through promotion so that he gets a greater profit.

There is two separate logins for the retailers and customers separately where when the customers login then they get to buy the items and when the retailer logins then he gets a different page where he can see the demand of the various items for the previous week and also the predicted value for the following week in numerical for all the items present in the store

The web page accessed by both the customers and the retailers is programmed with node.js, JSON and HTML languages. The prediction of the amount of goods to be ordered for the following week is done with Python 3.6. Prediction using Machine Learning can be done with various programming languages, namely Python, JSON, etc., but we have chosen Python as it is a high level interpreting programming language and it is easily readable by everyone. The prediction was accurate with an accuracy of approximately 75%. This depicts that it can be used in real-time and will surely help the retailers manage the buying and selling of their goods with a profit, by making all the goods to be sold out without any deterioration of the goods in the warehouse.

7. Advantages and Disadvantages

The advantages of this model are:

- ❖ The user interface is user friendly
- ❖ The machine learning prediction is highly reliable
- ❖ It prevents wastage of valuable perishable food products
- ❖ The retailers are benefited as their losses are reduced
- ❖ The customers get good quality food product as retailers buy food products as per their needss

Some of the disadvantages are:

- ❖ It cannot be implemented based on various seasons
- ❖ At the time of traffic in the cloud, the server may get slowed down

8. Applications

- ❖ This idea can be used to predict the future sale of a store based on its previous performances.
- ❖ It suggests the required amount of perishable food products to be bought and stored in the warehouse by the retailer for better sales.

9. Conclusion

This project not only concentrates for the benefit of retailers, but also customers. The retailers will know how much quantity of goods they must need to store in the warehouse and reduce the wastage of goods which in turn reduces their loss. At the same time, there will not be any shortage of goods to sell to the customers. The customers will also get good quality food products as the retailers buy the products as per the needs of the customers. This will also improve the customer-retailer relationship and will be a profit for both of them, in terms of money spent by on goods and quantity of food products. This model will enhance the sales of retailers because of good retailer-customer relationship. Thus, this project will help and benefit both the retailers and customers.

For more accuracy we may also use AI technology with Time Series Analysis which will use Weather Forecasting Algorithm to predict the next week purchase with the help of the current week. This will have an extension and depends on the parameters like geographical, season, etc. The same Random Forest Regressor Algorithm used in this project can be extended with season and geographical conditions if and only if there is data about them.

10. Future Scope

The project can still be extended to stores with branches based on the location of each branch and the prediction will depend on the respective areas too. Further, Market Basket Analysis can be implemented to analyze the customer's purchase patterns and to find the combination of products that are bought together by most of the customers. Also, recommender systems can be included to suggest recommendations to the customers.

11. Bibliography

- [1] Si-Yeong Lim, Sun Hur (2015) Determination of Optimal Shipping Quantity for Perishable Goods under Probabilistic Supply, Hindawi Publishing Corporation Mathematical Problems in Engineering, Vol. 2015, Article ID 274213
- [2] Sachin Khandekar, Aleksandra Titova (2012) Optimizing the Distribution Network of Perishable Products to Small Format Stores, Engineering Systems Division in Partial Fulfillment of the Requirements for the Degree of Master Engineering in Logistics.

- [3] Dilupa Nakandala, Henry Lau, jingjing Zhang (2016) Cost-Optimization Modeling for Fresh Food Quality and Transportation, *Industrial Management & Data Systems*, Vol. 116, No. 3, pp. 564-583
- [4] Wenzel, Hannah; Smit, Daniel; Sardesai, Saskia (2019) A Literature Review on Machine Learning in Supply Chain Management, *Artificial Intelligence and Digital Transformation in Supply Chain Management*, Vol. 27, ISBN 978-3-7502-4947-9, pp. 413-441
- [5] Jacob Huber, Aus Ludwigshafen am Rhein (2019) Data-driven Decision Support for perishable Goods
- [6] Reiner Jedermann, Mike Nicometo, Ismail Uysal, Walter Lang (2014) Reducing Food Losses by Intelligent Food Logistics, *Philos Trans A Math Phys Eng Sci.*, PMID: PMC4006167
- [7] Yared Lemma, Daniel Kitaw, Guletat Gatew (2014) Loss in Perishable Food Supply Chain: An Optimization Approach Literature Review, *International Journal of Scientific & Engineering Research*, Vol. 5, Issue 5, ISSN 2229-5518, pp. 302-311
- [8] Shuai Yang, Yujie Xiao, Yong-Hong Kuo (2017) The Supply Chain Design for Perishable Food with Stochastic Demand, *Sustainability* 2017,9,1195; doi: 10.3390/su9071195