

OPTIMIZED WAREHOUSE MANAGEMENT

ADITI TARATE
APURVA BASAPURE
SAGAR GHORPADE
(TEAM IGNITED NEURONS)

TABLE OF CONTENT

1. INTRODUCTION	3
1.1 OVERVIEW	3
1.2 PURPOSE	4
2. LITERATURE SURVEY	4
2.1 EXISTING PROBLEM	4
2.2 PROPOSED SOLUTION	5
3. THEORETICAL ANALYSIS	6
3.1 BLOCK DIAGRAM	7
3.2 HARDWARE/SOFTWARE DESIGNING	7
4. EXPERIMENTAL INVESTIGATIONS	10
5. FLOWCHARTS	11
6. RESULTS	11
7. ADVANTAGES & DISADVANTAGES	14
7.1 ADVANTAGES	14
7.2 DISADVANTAGES	15
8. APPLICATION	16
9. CONCLUSION	16
10. FUTURE SCOPE	16
11. BIBLIOGRAPHY	16
APPENDIX	17
SOURCE CODE	17

1. INTRODUCTION

Warehouses are the ever needing place to connect the local vendors and the people and companies that want to sell their products.

So a warehouse should be fully equipped to take in the products and take care so the products don't spoil till they reach the vendors. Also for a warehouse to keep making maximum profits and not letting any products waste is a priority.

So warehouse management is something that would bring additional costs if not properly looked up to, like it may cause over buying or under buying of goods, spoiling of goods due to climatic change, seasonal food issues causing expiring of products.

The main problem statement thus is:

A food delivery service has to deal with a lot of perishable raw materials which makes it all, the most important factor for such a company is to accurately forecast daily and weekly demand. Too much inventory in the warehouse means more risk of wastage, and not enough could lead to out-of-stocks - and push customers to seek solutions from your competitors. The replenishment of the majority of raw materials is done on a weekly basis and since the raw material is perishable, the procurement planning is of utmost importance,

And a simple solution that can be added to the warehouse management is implementing machine learning. Along with it smartly managing all the sensors with the help of a brain that is auto-adjusting and would find ways for itself to keep the warehouse working not compromising the actual profits.

1.1 OVERVIEW

This document will be focusing on the problems of a warehouse that delivers perishable food products and how it can be tackled using Machine Learning by making an app and web page interface to help track the demand and stock status.

Also, warehouses play a vital role in the storage and transportation of goods across the country. Hence it is of utmost importance that warehouses are maintained properly to ensure the quality of inventory. They play a major role in supply chain management. Warehouse Management Systems or WMS are gaining popularity as they provide ease of operation and increase efficiency. One such aspect of WMS is demand forecasting. In demand forecasting, historical sales data is collected and used to develop an estimate of consumer demand. Demand forecasting helps in decision making, calculating profit margins and also ensures minimum

wastage of inventory. The aim of this project is to predict the demand of the goods stored in The warehouse for the next 10 weeks or near future and send alerts about weather to ensure smooth operations inside and outside the warehouse. The historical sales data about the warehouse is collected which contains attributes like product code, product category, warehouse id, the date, current stock of the product, its demand for that day, and the stock margin. This data undergoes cleansing and wrangling to make it fit for use. Some redundant columns are removed from analysis. ARIMA model is used for forecasting and fairly accurate predictions are made. This forecasted data can be accessed through an android application which is built on the MIT app inventor.

1.2 PURPOSE

The purpose of this report is to maximize the profits of a warehouse by taking steps to perform better in the rising competition and avoid wastage of products due to any natural or seasonal causes.

This would reduce a ton of manual workload but would require a ton of prerequisites in terms of data for a start.

2. LITERATURE SURVEY

The two very promising sites for a jumpstart were:

1. <https://www.altexsoft.com/blog/demand-forecasting-methods-using-machine-learning/>

This link gives a simple way of defining why and how can we predict the demand for goods, in addition. Stating the facts that Machine Learning should be only implemented for volatile business platforms.

2. <https://www.kaggle.com/shashkhr25/food-demand-forecasting-challenge/data>

Again the most important parameter for machine learning is data. The more the data the more precise will be the predicted data. So for a start, we can infer how a data set is made and what all parameters need to fit into it.

2.1 EXISTING PROBLEM

The main work of a warehouse management team or a system is to

1. Keep track of present inventory
2. Provide condition to avoid food spillage
3. Manage inbound receipts and putaway.
4. Order new stock in accordance with the demand behavior

There are a series of problems with each of the task of warehouse management:

1. Keep track of the present inventory:

A warehouse would not necessarily get the products from the same companies so it will surely deal with a lot of different product codes. Still, traditional barcodes are being used which carry a very small amount of data which is also codes into bits which is not readable by the people.

2. Provide condition to avoid food spillage:

Changes in the surrounding condition of particular perishable goods may not always be noted by the manual warehouse management system.

3. Manage inbound receipts and putaway:

The billing process is done mostly on papers and by hand at most of the places increasing the chances of having manual error in the accounts due to the many unsold products and other major parameters.

4. Order new stock in accordance with the demand behavior:

As there are periods when there are heavy fluctuations in demand of particular products like functions, disasters, seasonal changes, etc.

2.2 PROPOSED SOLUTION

A whole new package can be put into work which will help in keeping transparency, making better predictions and easy tracking of packages.

1. RFID Tags
2. App development
3. Machine learning
4. IoT System.

All the problems faced can be tackled at a very small cost which will help to increase the revenue by reducing manual work practices, spillage of perishable goods

1. RFID Tags

Generally, a product comes with a datasheet which needs to be provided by the seller. But if the use of RFID tags is brought into picture it all eliminates, it will be helpful in the inventory management will help to provide the updated and changed datasheets of the products.

2. App development

To have remote access and an update on the stocks and whereabouts, an app can be developed for the staff for saving time and knowing the status of the warehouse.

3. Machine learning

Machine Learning will be used to build a time-series model. In this project, we used the ARIMA model for predicting the demand and it gave fairly accurate predictions. A time series is a sequence where a metric is recorded over regular time intervals. In time-series forecasting current data points depend on the past data points

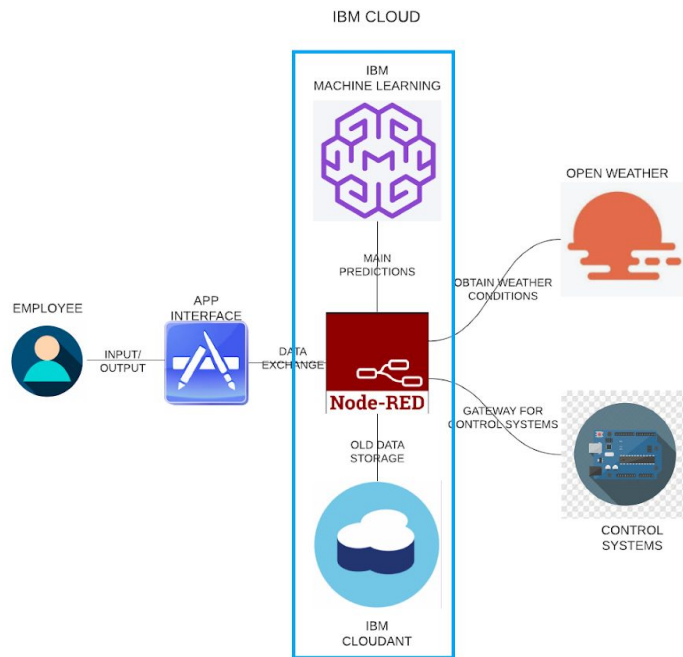
4. IoT System

Some other concerns are cost-cutting on electricity, security manual errors in operating cold storages, etc. This can be tackled by using a few sensors and controlling it with the help of a single server on which the data of RFID will be stored along with the predictions.

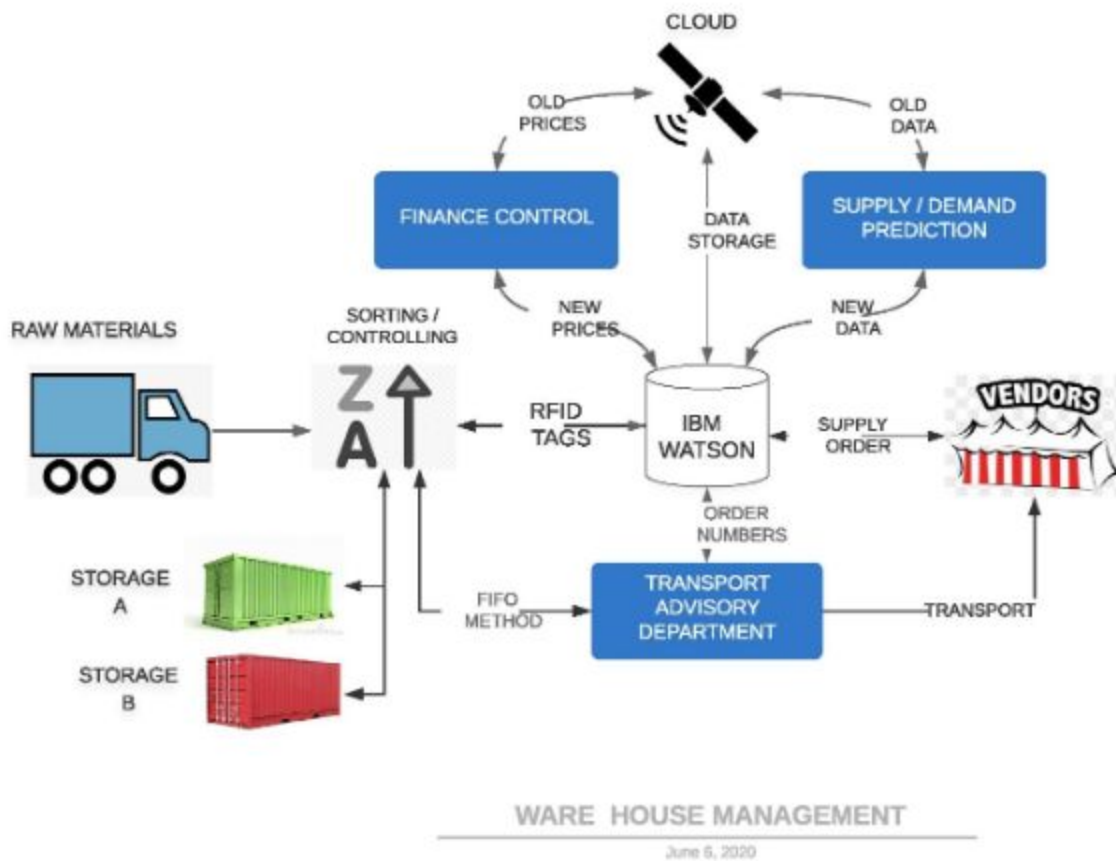
3. THEORETICAL ANALYSIS

How did we exactly plan to do it:

- **Step 1**: We started with the selection of a Dataset referring to our problem statement.
- **Step 2**: Then with the help of machine learning services provided by IBM Cloud and ARIMA model we started predicting values by feeding 1 year of sales data to the model.
- **Step 3**: Side by side using the IOT services by IBM Cloud we linked the sensors and data obtained by the open weather program.
- **Step 4**: Using the Flask framework we deployed the output by the Arima machine learning model and also the data from the IoT services using node-red as a gateway.
- **Step 5**: Finally develop an app using MIT App Inventor services linking all the services operated by the employees



3.1 BLOCK DIAGRAM



3.2 HARDWARE/SOFTWARE DESIGNING

Softwares used:

1. IBM cloud services
2. MIT app inventor
3. Python
4. Jupyter notebook
5. IBM Watson IoT Platform

The softwares used in the project can be divided into 2 main categories:

- Machine learning model
- IoT Project

The **Machine Learning Model** is divided into:

1. **Importing the dataset**

The dataset contains the data to be analyzed. We created 5 different datasets each for one product. The products were: dairy, groceries, oil, vegetables, and meat. The dataset set contained time-series data of 1 year, from 1-07-2019 to 30-06-2020, representing the demand of each date.

2. Data Wrangling

The date column which was a string object was converted to the index column as it is unique for each value and helps to analyze the data. Also some redundant columns

3. Choose the model for analysis

We experimented with a series of models to determine the best model to choose for analysis.

AutoRegressive Integrated Moving Average (ARIMA) model was selected to proceed with the analysis. ARIMA is actually a class of models that 'explains' a given time series based on its own past values, that is, its own lags and the lagged forecast errors, so that equation can be used to forecast future values. Any 'non-seasonal' time series that exhibits patterns and is not a random white noise can be modeled with ARIMA models. ARIMA model uses 3 terms as parameters to be passed while forecasting.

- p: The order of the AR term.
- d: The degree of differencing.
- q: The order of the MA term.

p refers to the number of lags that are actually useful for predictions. d refers to how many times do we have to difference the data to get stationary series. A stationary series is one that has constant mean and variance throughout different time periods. q refers to the number of lagged forecast errors we have to take into account.

A pure **Auto-Regressive (AR only)** model is one where Y_t depends only on its own lags. That is, Y_t is a function of the 'lags of Y_t '.

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + \epsilon_t$$

Likewise a pure **Moving Average (MA only)** model is one where Y_t depends only on the lagged forecast errors.

$$Y_t = \alpha + \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + \dots + \phi_q \epsilon_{t-q}$$

An ARIMA model is one where the time series was differenced at least once to make it stationary and you combine the AR and the MA terms.

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + \dots + \phi_q \epsilon_{t-q}$$

This model produced better results than other models.

- **Cloudant node:** Stores sensor data in database.
- **Delay node:** Limits the rate of messages to 1 message per 4 hours.

4. Train algorithm

The algorithm was trained on 11 months of data from June 2019 to May 2020. It used the forecast method to forecast the data. The p, d and q values were calculated by Auto ARIMA method in which the minimum and maximum values of p, d and q were given as parameters and the best value of p,d and q was obtained. The best value is one which has the minimum AIC. The Akaike Information Criteria (AIC) is a widely used measure of a statistical model. It basically quantifies 1) the goodness of fit, and 2) the simplicity of the model into a single statistic.

5. Test algorithm

The model was tested with the test data consisting of 30 values of June 2020 and it made predictions with different r2 scores. The p, d and q values for each product were:

Dairy: (17, 1, 1)

Groceries: (25, 1, 1)

Oil: (25, 1, 1)

Veggies: (18, 1, 1)


Meat: (18, 1, 1)

Deploying the model

6. App

For the user interface of the project we have developed an application using MIT app inventor. Through the app we have accepted the input in the form of date and product and subsequently predict the demand throughout the model. MIT app inventor provides a good GUI approach for app development in which we can grab and drop whichever component required for our App UI. For the interaction of the app it provides blocks to develop the back end in which the components can interact with each other and give the required output. Through URL and json files we can also interact with the internet.

Predict the Tomorrow Today!



The image shows a hand in a suit sleeve pointing at a target on a line graph. The graph has a solid orange line and a dashed grey line with circular markers. Below the graph are two blue buttons: 'Select date' and 'Select Product'. Under 'Select date' is a text input field with 'Date' as a label and '15th August, 2020' as the value. Under 'Select Product' is a text input field with 'Product' as a label and 'Groceries' as the value. At the bottom is a large green button labeled 'Predict Demand'.

Select date	Select Product
Date	15th August, 2020
Product	Groceries

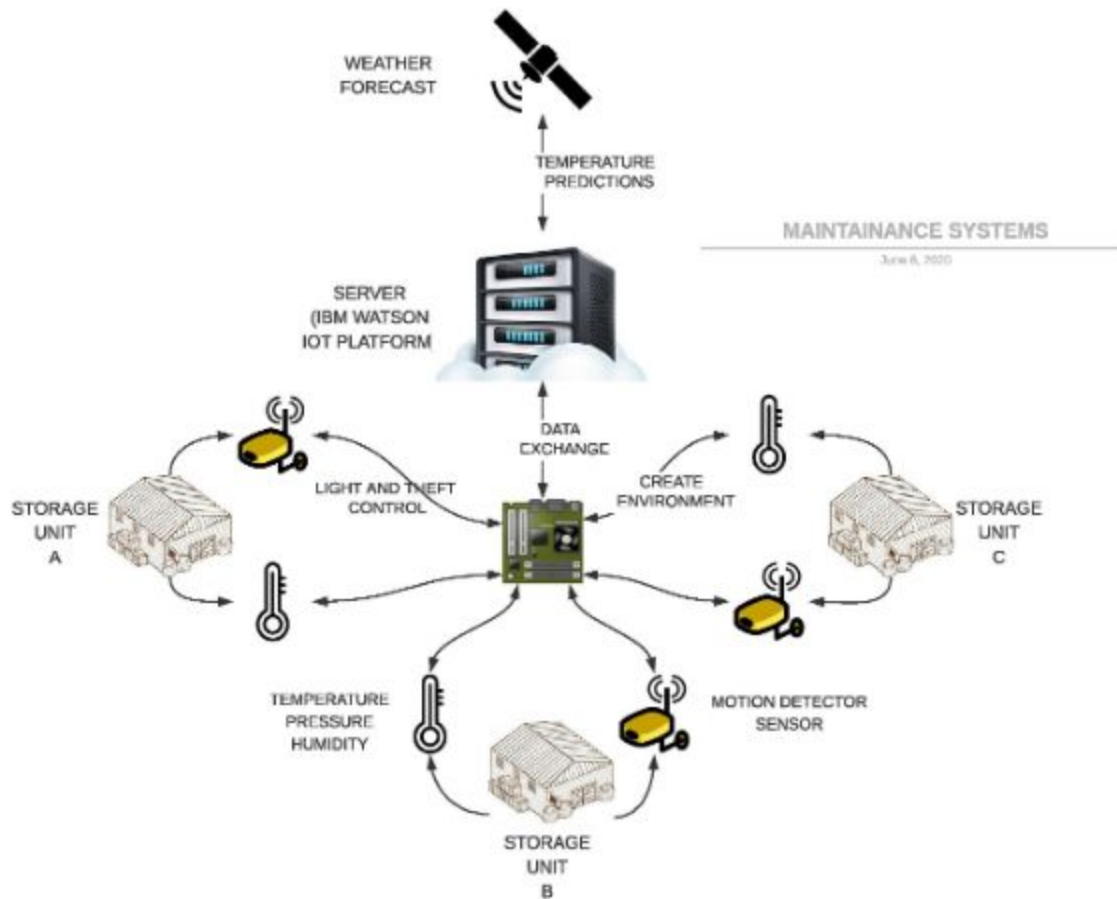
Predict Demand

4. EXPERIMENTAL INVESTIGATIONS

We experimented with a series of models for predicting demand.

- **Linear Regression:** It simply doesn't work well with time-series data. It fails to understand the relationship between past values and current values, which is the case with time-series data. When tested it gives the average of all the values as a constant graph.
- **Random Forest Regression:** For any data that a random forest hasn't seen before, it can predict an average of training values that it has seen before. This is not particularly useful, especially when we have an increasing or decreasing trend or a pattern in the data. Therefore, the Random Forest model doesn't go with time-series data and gives output the same as a Linear Regression model.
- **AR model:** Auto Regressive model gave pretty good results, recognizing the pattern to some extent and it was an improvement from Linear Regression and Random Forest. But still it wasn't enough. More precision was needed that is why the fourth and final ARIMA model came into the picture.
- **ARIMA model:** It showed the best results relative to all the 4 models. Although by varying the test data r^2 values vary a bit and it gives best results when the test data is less.

5. FLOWCHARTS



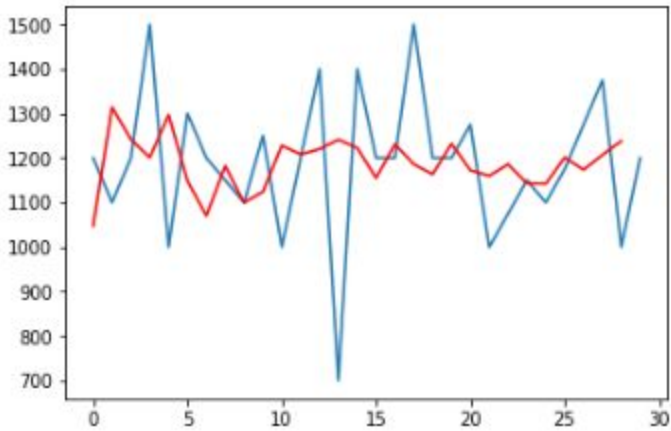
So to give an idea of how things will be working, all the sensors will be connected to the raspberry pi board which will be connected to the internet transmitting the data which will be stored into the IBM cloud servers. This and the data obtained by the open weather will be combined together making an environment to send alerts about harsh weather conditions.

6. RESULTS

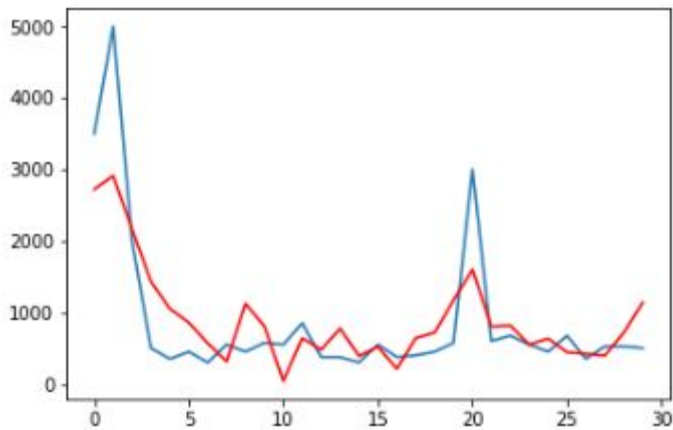
We used a series of different machine learning models, some of them which gave us a good success rate were the ARIMA, SARIMA, AR, etc. after comparing the complexity of all the models we decided to use the ARIMA model which is an abbreviation of **AutoRegressive Integrated Moving Average**. With the help of this model, we were able to achieve accuracies of a maximum of 93% for a dataset and a minimum of 24%.

The following are the graphs that were obtained after the dataset of 5 different food items that have different demand-type throughout the year. The red lines are the predicted supply while the blue lines are the actual demand values.

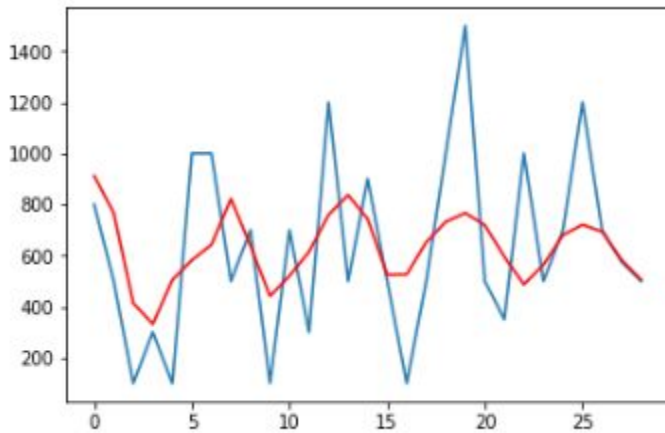
dairy: accuracy 22%



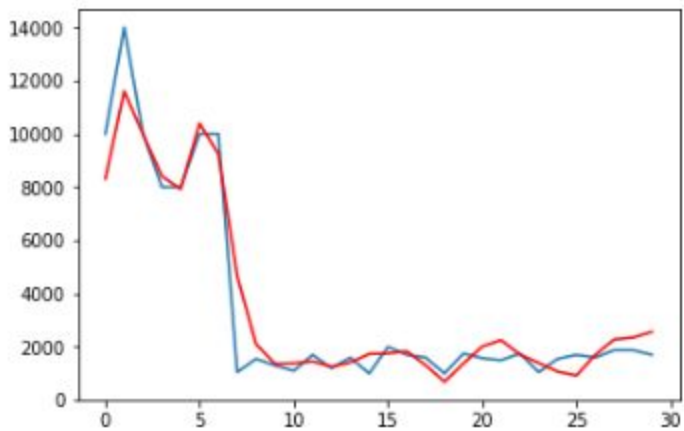
Oil: accuracy 68%



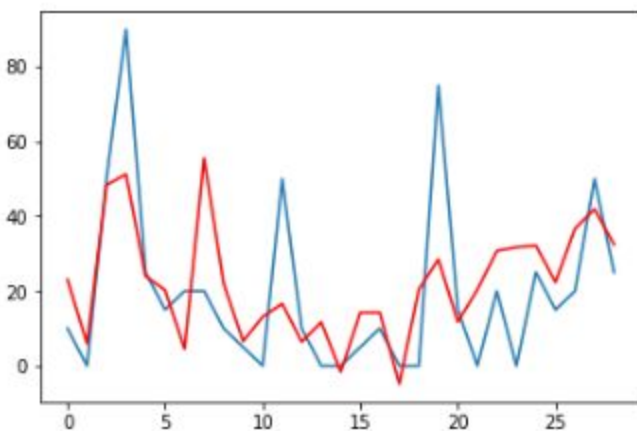
Veggies: accuracy 34%



Groceries: accuracy 93%

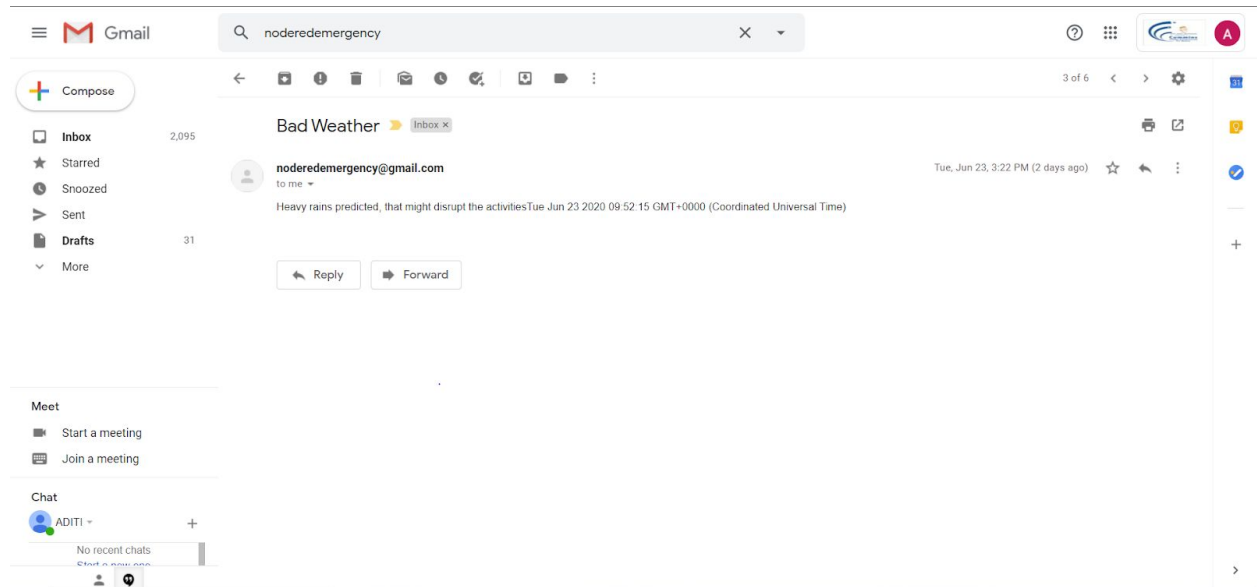


Meat: accuracy 45%



IoT Project:

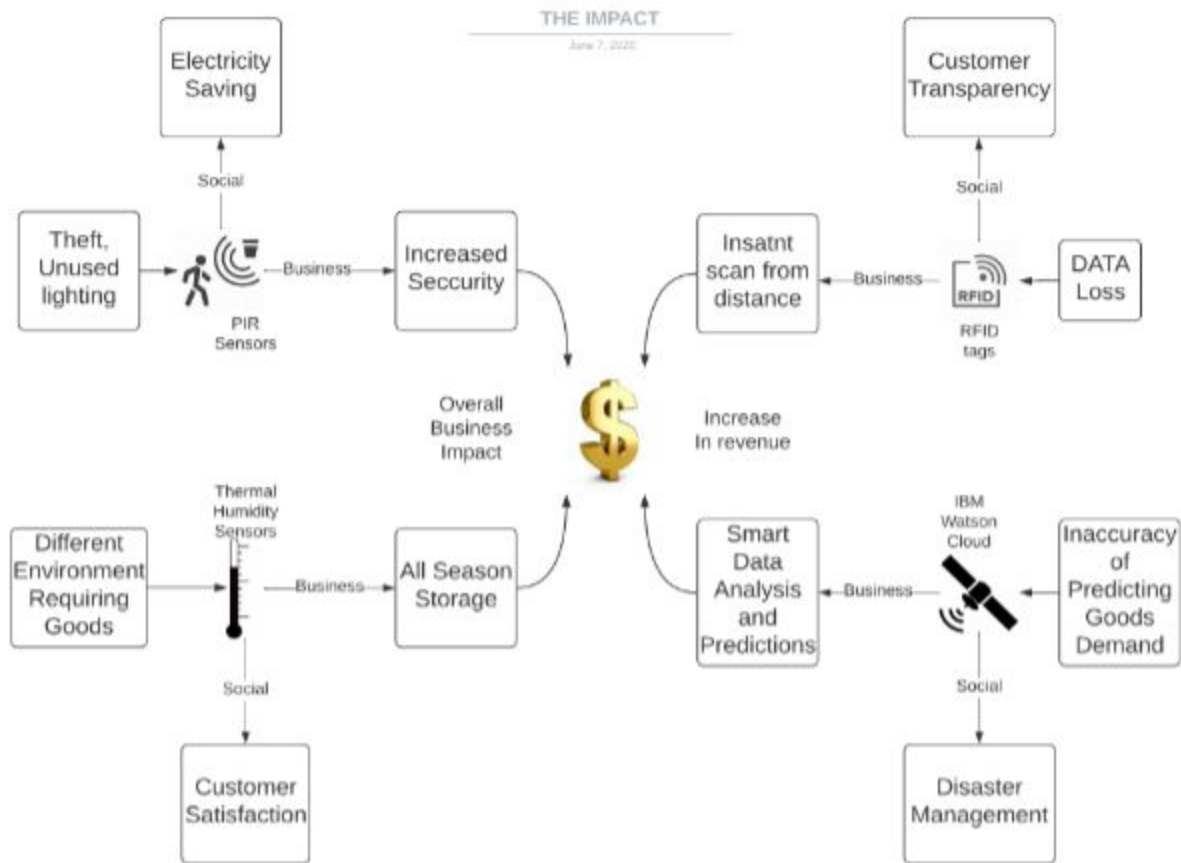
Alerts were sent on email whenever the weather became bad or the temperature of the warehouse increased beyond a certain threshold.



7. ADVANTAGES & DISADVANTAGES

7.1 ADVANTAGES

1. **Easily identifies trends and patterns:** Machine Learning can review large volumes of data and discover specific trends and patterns that would not be apparent to humans.
2. **No human intervention needed (automation):** With ML, you don't need to keep an eye on all the parts manually every day thus reducing the manual errors. Since it means giving machines the ability to learn, it lets them make predictions and also improve the algorithms on their own.
3. **Better Inventory Management:** Better sorting as more information about a particular product will be stored into the RFID tags of the particular products helping in grouping various products together based on their storing conditions.
4. **Getting updates right on phones:** The employees will get direct updates on their phones which will help them give the required human touch to the inventory management



7.2 DISADVANTAGES

1. **Initial cost of equipment:** Automated equipment is more expensive up front than manual equipment. However, what you save in man power and what you gain in increased productivity, means the equipment will eventually pay for itself and then some.
2. **Reduced flexibility for change:** Once automated systems are in place, it is likely not as easy to make changes in your workspace. But once you go automated and see how smoothly everything runs, it's not likely you'll want to return to manual equipment afterward.
3. **Possible downtime due to malfunction:** With automatic machines there is always the chance of a problem or breakdown, which can lead to considerable downtime while it is repaired. If the problem cannot be fixed by anyone on site, an outside specialist may need to be called, which could mean more time spent waiting. In some cases, work may be able to continue manually in the meantime, but in some cases that isn't possible. Equipment malfunction can be avoided in most cases with routine maintenance of all

machinery. If proper care is taken to keep automated systems in good shape, breakdowns should only occur rarely, if ever.

4. **Maintenance costs:** Some automated equipment needs maintenance. Routine maintenance may be performed regularly by onsite workers, but periodic professional maintenance should be handled by specialists who are trained to inspect and tune-up machinery so that it runs smoothly and efficiently. Properly maintained equipment will save you money in the long run by preventing problems and increasing the overall lifespan of your systems. Carolina Material Handling offers many types of automated equipment that require very low maintenance, keeping costs low.

8. APPLICATIONS

- This project will have wide applications in supply chain management and retail sales.
- Faster delivery systems yield increase in the in and out of goods speed.
- Followed by using connected devices, warehouse workers can instantly identify products and packages.

9. CONCLUSION

- Grading on neutral grounds, in financial terms the advantages outweigh the maintenance and initial costs heavily.
- Considering the increasing competition and increasing population, having a variety of products is necessary which brings the need of various new hardware necessities, whose control with the help of automation is the best option.
- Relying heavily on Machine Learning is not advisable one should always keep an eye on market trends and be prepared for disastrous situations like COVID 19.

10. FUTURE SCOPE

We have the forecasted data of about 1 year which was forecasted from previous year's data. The steps that can be taken in due future to enhance the project can be collecting the real-time data of everyday demand and retrain the model everyday with that data to increase the accuracy of predictions.

Also we can start working on actual sensors and connect them to node red via a gateway for monitoring the warehouse environment.

This also raises questions about the accuracy of the forecasts and how it can be improved. This can be done by refining the data more and some more data about the seasonality of the products.

11. BIBLIOGRAPHY

- Kaggle Datasets

<https://www.kaggle.com/datasets?sortBy=relevance&group=public&search=Retail+&page=1&pageSize=20&size=all&filetype=all&license=all>

- Demand forecasting

<https://blog.arkieva.com/demand-forecasting/#:~:text=Demand%20Forecasting%20is%20the%20process,purchase%20in%20the%20foreseeable%20future.>

- ARIMA model

<https://www.machinelearningplus.com/time-series/arma-model-time-series-forecasting-python/>
<https://towardsdatascience.com/analyzing-time-series-data-in-pandas-be3887fdd621>
https://www.statsmodels.org/dev/_modules/statsmodels/tsa/arma_model.html

- Linear Regression

<https://towardsdatascience.com/a-beginners-guide-to-linear-regression-in-python-with-scikit-learn-83a8f7ae2b4f>

- Random Forest Regression

<https://towardsdatascience.com/random-forest-and-its-implementation-71824ced454f>

- Building ARIMA model

<https://medium.com/ibm-watson/leverage-watson-machine-learning-function-deployments-in-data-science-workflows-476ce9ed726a>
<https://dataplatform.cloud.ibm.com/exchange/public/entry/view/a8819ce0e589432663a01f38a4980075>

APPENDIX

A. SOURCE CODE

```
[{"id":"644241a7.bde88","type":"tab","label":"IoT
project","disabled":false,"info":""},{id:"95d923c6.30d43","type":"ibmiot
in","z":"644241a7.bde88","authentication":"apiKey","apiKey":"31f52016.25558","inputType":"evt"
,"logicalInterface":"","ruleId":"","deviceId":"Sensor","applicationId":"","deviceType":"Sensor_data"
,"eventType":"","commandType":"","format":"json","name":"IBM
IoT","service":"registered","allDevices":"","allApplications":"","allDeviceTypes":false,"allLogicalInt
erfaces":"","allEvents":true,"allCommands":"","allFormats":"","qos":0,"x":90,"y":160,"wires":[["8ae
```

```

61e74.c5893","3bafa870.eed9a8","f2014723.903288","c6422a95.d31db8"]]],{"id":"e293dc09.5d
27b","type":"function","z":"644241a7.bde88","name":"Globally set the
variables","func":"global.set(\"WarehouseTemperature\",msg.payload.d.temperature)\nglobal.set
(\"Humidity\",msg.payload.d.humidity)\nglobal.set(\"ColdStorageTemperature\",msg.payload.d.o
bjectTemp)\nreturn
msg;","outputs":1,"noerr":0,"x":490,"y":100,"wires":[["325e5de0.7ba972"]]],{"id":"713682fd.d5d4a
c","type":"cloudant
out","z":"644241a7.bde88","name":"","cloudant":"","database":"sensor_data","service":"node-red
-jecca-cloudant-1591687694923-2210","payonly":false,"operation":"insert","x":590,"y":140,"wires
":[[]],{"id":"5a99da4b.200054","type":"function","z":"644241a7.bde88","name":"","func":"
msg.payload = {\n  _id:msg.payload.d.temperature,\n  ColdStorage_temp :
msg.payload.d.temperature,\n  Humidity : msg.payload.d.humidity ,\n  Warehouse_temp :
msg.payload.d.objectTemp,\n  }\n}\nreturn
msg;","outputs":1,"noerr":0,"x":430,"y":140,"wires":[["713682fd.d5d4ac"]]],{"id":"325e5de0.7ba97
2","type":"debug","z":"644241a7.bde88","name":"","active":true,"tosidebar":true,"console":false,"t
ostatus":false,"complete":"payload","targetType":"msg","x":710,"y":100,"wires":[[]],{"id":"a874c17
a.7fc46","type":"function","z":"644241a7.bde88","name":"warehouse temp","func":"msg.payload
= msg.payload.d.objectTemp\nreturn
msg;","outputs":1,"noerr":0,"x":480,"y":180,"wires":[["feee1551.c33778"]]],{"id":"3a03437a.4b37b
c","type":"function","z":"644241a7.bde88","name":"Build Mail","func":"msg = {\n  payload :
\"Temperature is too high, cooling process started.\" + Date().toString(),\n  topic : \"Warehouse
Temperature\",}\n}\nreturn
msg;","outputs":1,"noerr":0,"x":520,"y":420,"wires":[["92f71a53.fd6328"]]],{"id":"feee1551.c33778
","type":"switch","z":"644241a7.bde88","name":"Comparing Warehouse
temperature","property":"payload","propertyType":"msg","rules":[{"t":"btwn","v":"35","vt":"num","v
2":"60","v2t":"num"}],"checkall":"true","repair":false,"outputs":1,"x":740,"y":180,"wires":[["3a03437
a.4b37bc"]]],{"id":"92f71a53.fd6328","type":"e-mail","z":"644241a7.bde88","server":"smtp.gmail.c
om","port":"465","secure":true,"tls":true,"name":"aditi.tarate@cumminscollge.in","dname":"","x":
810,"y":440,"wires":[[]],{"id":"8ae61e74.c5893","type":"delay","z":"644241a7.bde88","name":"","p
auseType":"rate","timeout":"5","timeoutUnits":"seconds","rate":"1","nbRateUnits":"4","rateUnits":"
hour","randomFirst":"1","randomLast":"5","randomUnits":"seconds","drop":true,"x":260,"y":180,"
wires":[["a874c17a.7fc46"]]],{"id":"3bafa870.eed9a8","type":"delay","z":"644241a7.bde88","name
":"","pauseType":"rate","timeout":"5","timeoutUnits":"seconds","rate":"1","nbRateUnits":"4","rateU
nits":"hour","randomFirst":"1","randomLast":"5","randomUnits":"seconds","drop":true,"x":260,"y":
140,"wires":[["5a99da4b.200054"]]],{"id":"f2014723.903288","type":"delay","z":"644241a7.bde88
","name":"","pauseType":"rate","timeout":"5","timeoutUnits":"seconds","rate":"1","nbRateUnits":"4
","rateUnits":"hour","randomFirst":"1","randomLast":"5","randomUnits":"seconds","drop":true,"x":
260,"y":100,"wires":[["e293dc09.5d27b"]]],{"id":"2aab5dad.4219d2","type":"openweathermap
in","z":"644241a7.bde88","name":"","wtype":"current","lon":"","lat":"","city":"Pune","country":"India

```



```
":"644241a7.bde88","name":"Getting values","func":"msg.payload = {\n
  \"WarehouseTemperature\": global.get(\"WarehouseTemperature\"),\n
  \"ColdStorageTemperature\": global.get(\"ColdStorageTemperature\")\n}\nreturn
msg;","outputs":1,"noerr":0,"x":360,"y":580,"wires":[["75c5dca7.917714"]]},{"id":"75c5dca7.917714","type":"http
response","z":"644241a7.bde88","name":"","statusCode":"","headers":{},"x":550,"y":620,"wires":[]
},{ "id":"31f52016.25558","type":"ibmiot","z":"","name":"","keepalive":"60","serverName":"wqlv9j.m
essaging.internetofthings.ibmcloud.com","cleansession":true,"apld":"","shared":false}}
```

