CROP DEMAND PREDICTION SYSTEM

By Team:- Cheekati Rajya Lakshmi

Team Members :

1. Cheekati Rajya Lakshmi
2. Sheetal Kumari
3. R.Manikanata Krishna Charan
4. Jose James

PROBLEM STATEMENT:

Agriculture is one of important sectors of Indian economy. In every sector of production, there is a demand of those respective product types. Similarly in agriculture sector there are different demands for different crops. It is evident that there is a huge gap between demand and supply of various crops, due to which both farmers and consumers are facing problems. So latest technologies like data analytics, machine learning can be used to predict the product/crop in demand. The available data about the demand, supply, price variation of the crops and other factors affecting the supply chain of agricultural produce can be used to analyze and come up with a model to predict and forecast market variations of agricultural crops.

MARKET/CUSTOMER/BUSINESS NEED ASSESSMENT:

A business model can be drawn from the above-mentioned project idea. One can develop a business on the above ideas such that one can predict which crop will be in demand for next years, so that the person can encourage farmers to cultivate respective crop which will be in demand for next years. So that farmers can decide which crop he/she can cultivate according to their area, soil, weather, other conditions.

TARGET SPECIFICATIONS AND CHARACTERIZATION:

Our main aim/target customer is farmer. When a season of crop production starts, farmer can be in a dilemma of which crop to cultivate. When the machine learning model predict the crop in demand for the next years, the price of the crop will also be higher compared to the other crops. So, the farmer can earn much more compared to earnings through present crop. When a farmer has a detailed clarity about which crop is in demand, issues about the prices it provides workforce for rural workers also.

CONCEPT DEVELOPMENT:

The main way of developing this idea into reality is by using Machine Learning techniques. We have many algorithms/techniques of developing a trained machine learning model for purpose of predictions. Through the trained model, we predict the crop in demand. According to the forecast results, area of land, soil conditions etc we visualize the crop predictions. From this farmer can decide on which crop to be cultivated. In this concept development, we have many phases. They

are Data Collection, data pre-processing, machine learning algorithm, predicting the demanding crops and then the last step is visualizing. The critical farming decisions are made by analysing the price data. The price chart for TOP crops are prepared by gathering, reviewing and analysing the various data collected from the different authentic sources. This analysis helps the farmers in evaluating future demands of the crops. This report will help farmers determine the variety and the time of planting the crops.

FINAL PRODUCT PROTOTYPE WITH SCHEMATIC DIAGRAM:

There are totally 5 stages in the product development:

1. Data Collection
2. Data pre-processing
3. Model selection
4. Demand prediction
5. Results and visualization

1. Data Collection:

For the prediction of crop in demand for the future years, the main important aspect is data. Here we consider the data to be continuous i.e., there should be accurate data of each, and every year's crops produced, types of crops produced, nature of soil, overall crop consumption, quantity of the crop produced, how much quantity is sold out, what is the preserved quantity of crop, which crop is minimally available for the future years, so on. Data is gathered from authentic websites like Ministry of Agriculture & Farmers Welfare, Food and Agriculture Organization, APEDA, NITI Aayog.

2.Data pre-processing:

For training of machine learning models, pre-processing of the data is one of the main tasks for the accurate prediction of the results. As crop prices depends on various factors, the obtained data contains many outliers, null values etc. So, to avoid these conditions, we must pre-process the data. If we have outliers in our model, we can use inter quartile range to remove the

outliers. We can fill the null values by replacing them with mean or median. Since demand data was unaccounted in any authentic websites, simulated data has been used for analysis. As the data for demand was simulated, the required accuracy is partially met, when actual data is available, accuracy of the model can be increased.

3. Model Selection:

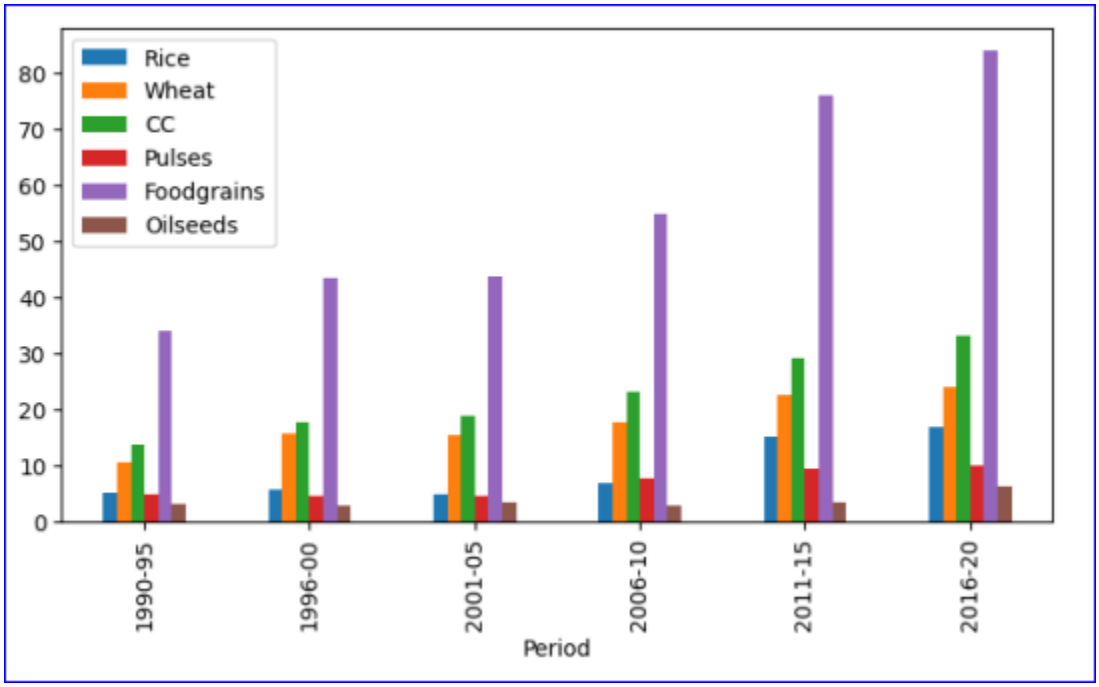
For any predictive system, the main aspect is the model which we want to train. In machine learning, technique selection is very important. We have many techniques in machine learning to tarin the model like supervised learning (Regression- linear regression, classification- logistic regression, decision trees, support vector machine, Random Forest), Unsupervised learning (clustering, Association) etc. For our crop demand prediction, we used regression. Our model uses regression because there are multiple crops data and our model have to predict the demand of those respective crops according to the previous year's statistics.

4. Demand prediction:

Go over past records. One of the most used indicators of current demand is past demand. Add up the total quantity of crop sold over the past year and pay attention to any seasonal trends that may be displayed by spikes or dips in the amount and quantity of product sold. The pitfalls with this method are that it doesn't account for changes in the marketplace. So, for this we use regression algorithms. Regression includes linear regression algorithm. Below is sample data of demand of each crop with their respected demands (in million tones): As below is a sample data of previous years demand of crops, The demand for the future years also can be near to the past years. Crop demand is dependent on various factors such as weather, soil, area of crop production, etc. So this is only sample data of previous years demand. Our data which we want to used for pre-processing should consist of various other factors. Then only our model can accurately predict the output.

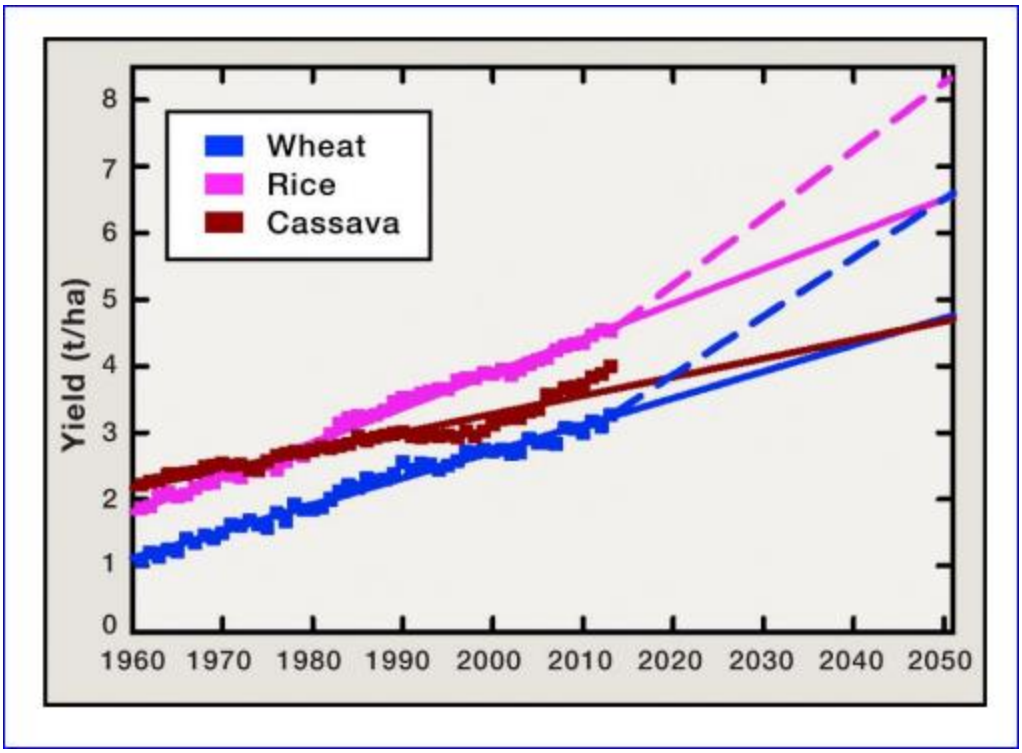
Period	Rice	Wheat	CC	Pulses	Foodgrains	Oilseeds
1990-95	5.16	10.54	13.65	4.68	34.03	3.04
1996-00	5.70	15.62	17.66	4.54	43.52	2.68
2001-05	4.89	15.28	18.91	4.55	43.63	3.46
2006-10	6.62	17.53	23.14	7.52	54.82	2.79
2011-15	15.12	22.45	29.06	9.36	75.99	3.32
2016-20	16.88	23.85	33.21	10.00	83.94	6.14

For the above data, the visual representation is shown below. For our sample data we took the crops Rice, Wheat, CC, Pulses, Foodgrains, Oilseeds. The data we have provides the visual representation of all the demands of previous years. So, from this past data we have to retrieve the future demand of the data.

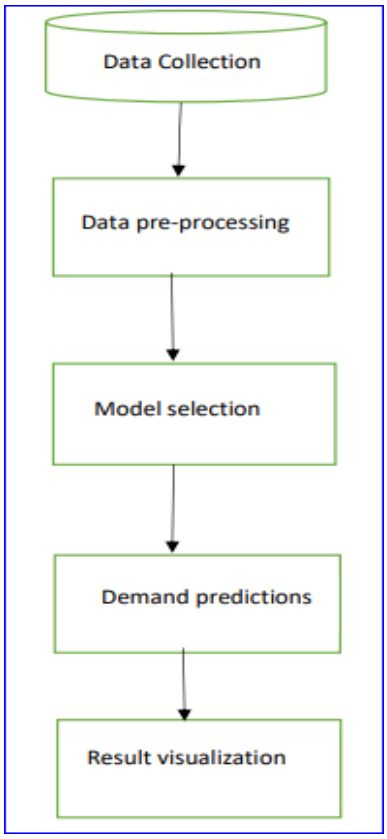


5. Results and Visualization:

Our model classifies which crop is in more demand compared to other crops. After the model training, we will predict which crop is in much demand. Our result deploys what is the demand of each crop for further years. After the model is trained, we have to test the accuracy of the model. If the accuracy of the model is good we can use the model and convert it into the business model. According to the results, we visualize the results by plotting them. Through these results we can decide which crop is in more demand according to the weather conditions, soil fertility, etc. Sample visualization of future predictions: Below visualization is a sample visualization of future/predicted results of the data using machine learning model. From the below visual representation we can say that we have the data of wheat, rice, Cassava's past years demand and past years statistics now we predicted the future demand of these crops.



Schematic diagram of implementation:



Model Implementation:

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
from sklearn import preprocessing, svm
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression

df = pd.read_excel("tomato_data.xlsx")
prices = []
for i in df["Today price"].tolist():
  price = i.split()
  prices.append(price[0])
supply = df["Today arrival"].tolist()
supply_lm = df["Last month arrival"].tolist()
supply_lw = df["Last week arrival"].tolist()

df = pd.read_excel("tomato_data.xlsx", "FEB 2020")
for i in df["Today price"].tolist():
  price = i.split()
  prices.append(price[0])
supply = supply + df["Today arrival"].tolist()
supply_lm = supply_lm + df["Last month arrival"].tolist()
supply_lw = supply_lw + df["Last week arrival"].tolist()

df = pd.read_excel("tomato_data.xlsx", "MAR 2020")
for i in df["Today price"].tolist():
  price = i.split()
  prices.append(price[0])

supply = supply + df["Today arrival"].tolist()
```

```
supply = supply + df["Today arrival"].tolist()
supply_lw = supply_lw + df["Last week arrival"].tolist()
supply_lm = supply_lm + df["Last month arrival"].tolist()
dates = list(range(31+29+31))

new_supply_lm = []
for i in range(len(supply_lm)):
  new_supply_lm.append(float(supply_lm[i])**3)

plt.scatter(new_supply_lm, prices)
plt.xlabel("Supply Quantity")
plt.ylabel("Price Demand")

regr = LinearRegression()

X = np.array(new_supply_lm).reshape(-1, 1)
Y = np.array(prices).reshape(-1, 1)

regr.fit(X, Y)
print(regr.score(X, Y))

y_pred = regr.predict(X)
plt.plot(new_supply_lm, y_pred, color ='k')

plt.scatter(new_supply_lm, prices, color ='b')
```

Results:

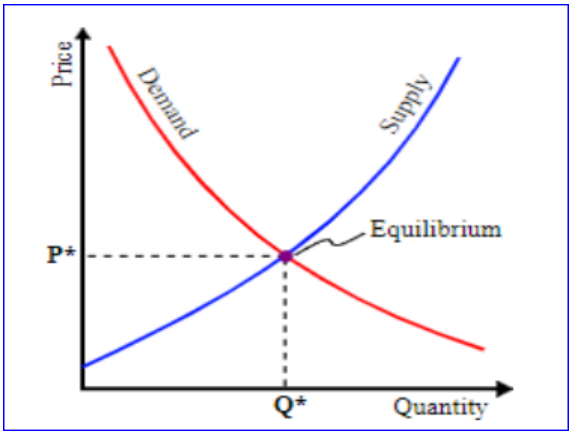


Fig-1 Demand, Supply and Price Inflation

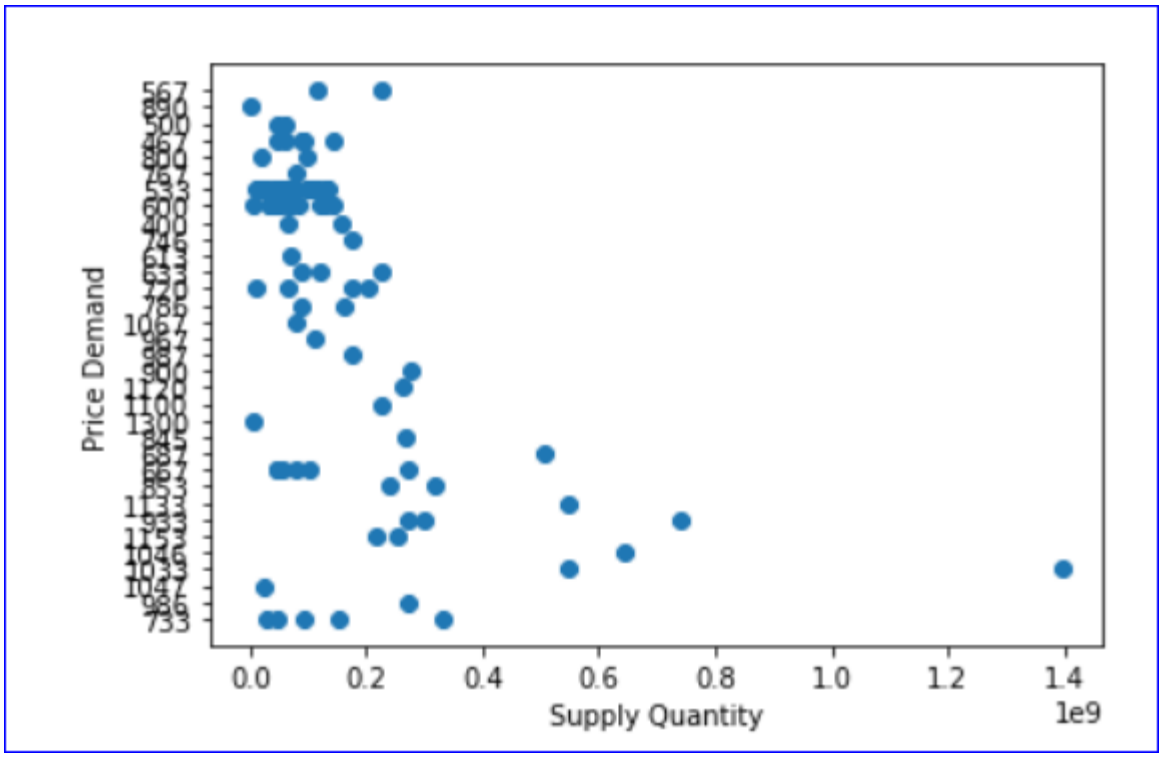


Fig -2 Plotting between Supply Quantity and Price Demand

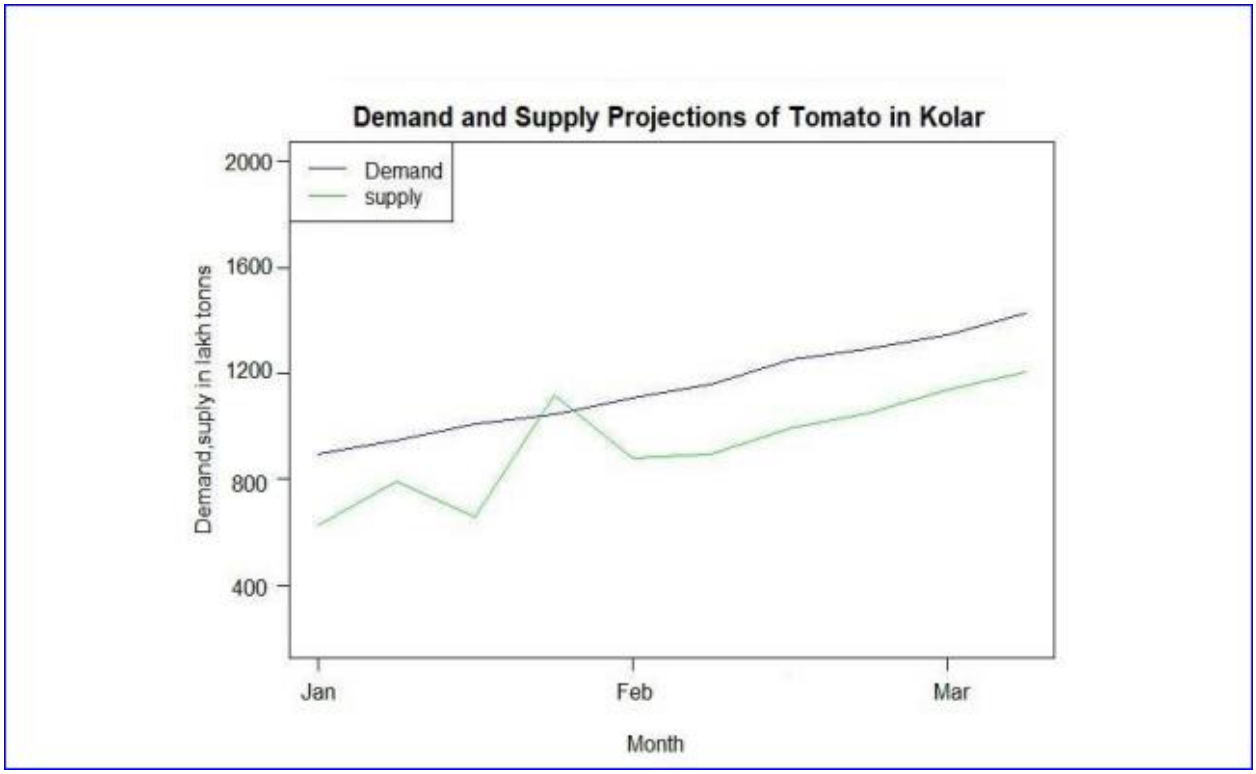


Fig -3 Demand Supply Projection for Tomato for the district Kolar from Jan to March

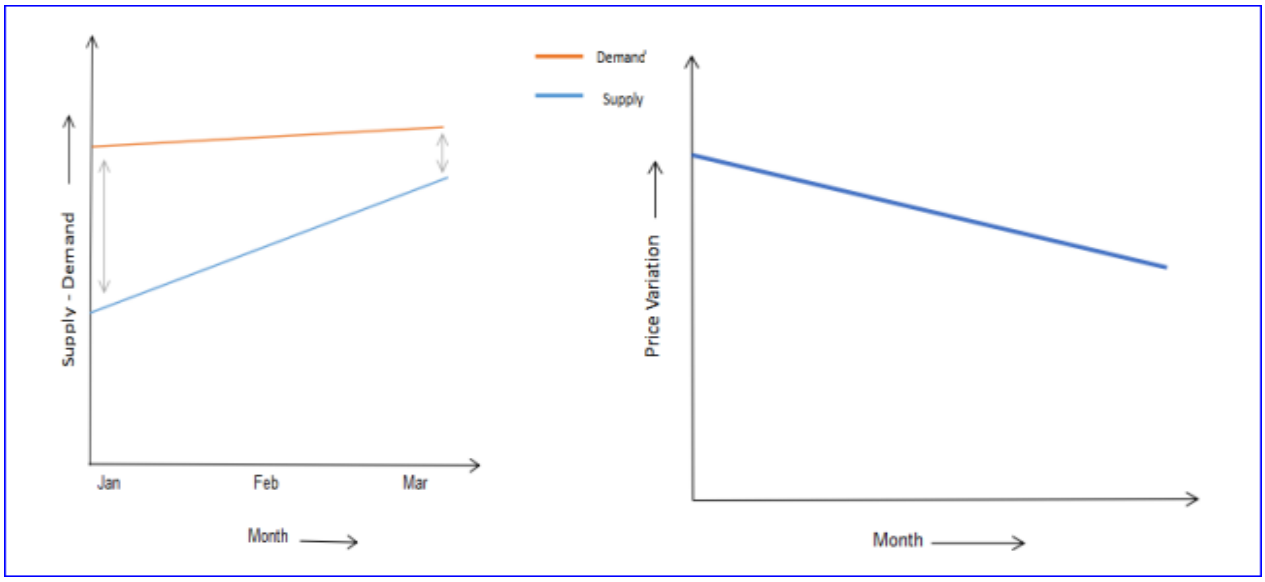


Fig -4 Reduced gap between demand &supply

Fig -5 Expected Price Prediction

CONCLUSION:

I hereby conclude that this is a business model idea which works on machine learning which is used to predict the future demanding crops. We can use regression to predict the crops in demand. We can use linear regression, support vector machine, decision trees etc algorithms to predict the demanding crop according to the weather conditions, soil fertility, water availability, past crop demand etc. By considering this dataset, a Supply-Demand Prediction forecasting model has been developed in this work that guides the farmers in selecting the appropriate crops to grow.

According to the findings, there is a significant mismatch between customer demand for different agricultural products and farmer supply of those same crops, which has caused market price fluctuations that are unanticipated and costly for both consumers and farmers. By assisting the system to direct farmers in choosing the appropriate crops to grow satisfying the actual needs of society (demand), this problem could be solved by the Demand-Prediction forecasting model developed in this work. This in turn satisfies the society's true need, avoiding losses for both farmers and customers at peak periods. The current issue can be effectively resolved by attaining equilibrium in the demand and supply of TOP crops with the help of this model.