Vegetable Price Prediction Using Machine Learning

GROUP MEMBERS

OMKAR S. MAHAJAN

PARTH P. SHIRKE

GAURAV S. SHUKLA

PROJECT GUIDE:

PROF. D.N. LONDHE

Abstract

- Just imagine! how much beneficial it will for us if we know the price of something in advance. Especially when it is the price of items which we use on daily basis.
- Vegetable is an indispensable part of every kitchen. We believe that, all will agree that it will definitely be profitable for every individual if we know price it in advance.
- Farmers can also get the benefit of knowing the price in advance.
- We have proposed a system which will predict the price of vegetable in advance using the concepts of Machine Learning.

Introduction

- Price of anything matter a lot in almost every aspect of life especially when we are taking about a middle-class family.
- More than 50 percent of total population fall under the category of middle and lower middle class(including population below poverty line). This is a huge figure, for them price of small things such as groceries, household items, vegetables, etc. matters a lot.
- Farmers are one of the most important stakeholder in agricultural sector, will also be benefited from this project.
- There are many factors which impact the price of vegetable such as climate, production cost, technological factors, storage possibilities, production, festive season, etc.

Literature Survey

Sr. No.	Title	Author	Publisher	YOP	Methodology	Limitations
1.	A Prediction Approach for Stock Market Volatility Based on Time Series Data	Sheikh Mohammad Idrees, M. Afshar Alam, Parul Agarwal	IEEE Access	2019	ARIMA model is used.	
2.	Vegetable Price Prediction Based of Particle Swarm Optimization – Back Propagation Neural Network	YE Lu, LI Yuping, LIANG Weihong	IEEE Conference	2015	BP – PSO model is used for prediction.	Implemented for single commodity
3.	Vegetable Price Prediction Using Data Mining Classification Technique	G. M. Nasira, N. Hemageetha	IEEE Conference	2012	The Back Propagation Neural Network Prediction Model is used.	Implemented for single market.

Objectives

- To enable common man to plan their budget strategies
- To help farmers in deciding and planning their crop plantation and harvesting related strategies

A Glance at Dataset

- Dataset has been collected from following government portals:

1. https://agmarknet.gov.in/

1. https://data.gov.in/

Dataset

Sl no.	District Name	Market Name	Commodity	Variety	Grade	Min Price (Rs./Quintal)	Max Price (Rs./Quintal)	Modal Price (Rs./Quintal)	Price Date
1	Ahmednagar	Ahmednagar	Potato	Other	FAQ	1000	2000	1500	22 Dec 2019
2	Ahmednagar	Ahmednagar	Potato	Other	FAQ	800	1600	1200	15 Dec 2019
3	Ahmednagar	Ahmednagar	Potato	Other	FAQ	1000	2200	1600	02 Dec 2019
4	Ahmednagar	Ahmednagar	Potato	Other	FAQ	1000	2200	1600	01 Dec 2019
5	Amarawati	Amrawati(Frui & Veg. Market)	Potato	Other	FAQ	1600	1800	1700	25 Dec 2019
6	Amarawati	Amrawati(Frui & Veg. Market)	Potato	Other	FAQ	1600	1800	1700	24 Dec 2019
7	Amarawati	Amrawati(Frui & Veg. Market)	Potato	Other	FAQ	1600	1800	1700	23 Dec 2019
8	Amarawati	Amrawati(Frui & Veg. Market)	Potato	Other	FAQ	1500	1800	1650	22 Dec 2019
9	Amarawati	Amrawati(Frui & Veg. Market)	Potato	Other	FAQ	1600	2000	1800	21 Dec 2019
10	Amarawati	Amrawati(Frui & Veg. Market)	Potato	Other	FAQ	1600	2000	1800	20 Dec 2019
11	Amarawati	Amrawati(Frui & Veg. Market)	Potato	Other	FAQ	1800	2000	1900	18 Dec 2019
12	Amarawati	Amrawati(Frui & Veg. Market)	Potato	Other	FAQ	1800	2200	2000	17 Dec 2019
13	Amarawati	Amrawati(Frui & Veg. Market)	Potato	Other	FAQ	1600	2000	1800	16 Dec 2019
14	Amarawati	Amrawati(Frui & Veg. Market)	Potato	Other	FAQ	1800	2000	1900	15 Dec 2019
15	Amarawati	Amrawati(Frui & Veg. Market)	Potato	Other	FAQ	1800	2000	1900	14 Dec 2019

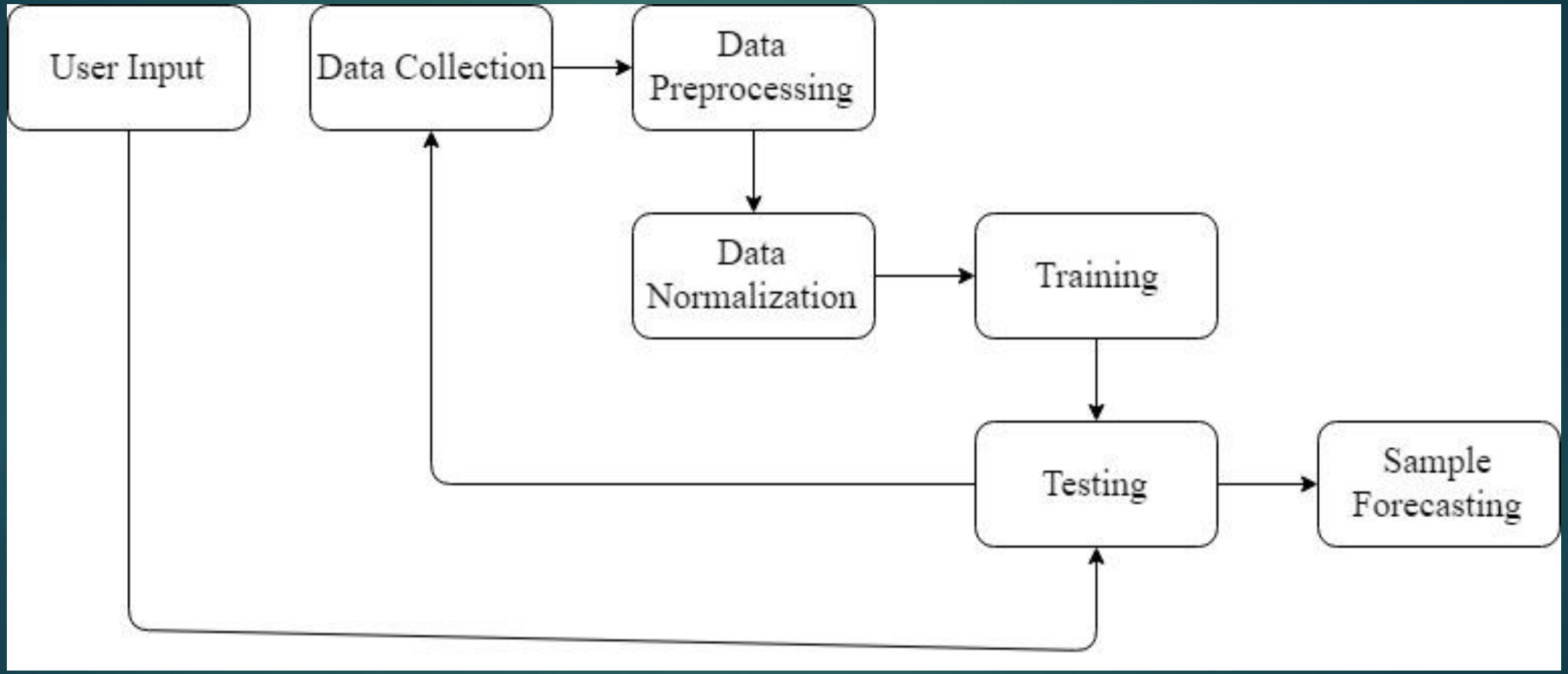
Snapshot of Dataset Collected from agmarknet for Potato Commodity for all over India

Parameters of Dataset

Following parameters are included in dataset collected from:

1. State
2. District
3. Market
4. Commodity
5. Variety
6. Min Price/Quintal
7. Max Price/Quintal
8. Modal Price/quintal
9. Arrival Date

Flowchart


User Input
Data Collection
Data Preprocessing
Data Normalization
Training
Testing
Sample Forecasting

Algorithm

Random Forest

The random forest is a supervised learning algorithm that randomly creates and merges multiple decision trees into one “forest.” The goal is not to rely on a single learning model, but rather a collection of decision models to improve accuracy.

Neural Network

Neural Network, is a computational learning system that uses a network of functions to understand and translate a data input of one form into a desired output, usually in another form. The neural network itself may be used as a piece in many different machine learning algorithms to process complex data inputs into a space that computers can understand.

Working of Algorithm

- **Following parameters has been considered while developing the system:**
  1. MinPrice/Quintal
  2. MaxPrice/Quintal
  3. ModalPrice/Quintal
  4. Rainfall
  5. Humidity
  6. Temperature
  7. Dew
  8. etc.....

- **All above parameter has been consider for training the model & making the model ready for prediction**

Working Steps of System

1. User register himself/herself with system.
2. User login into the system using credential .
3. User will be navigated to the prediction window.
4. A series of option will be provided to narrow down the burden on user.
5. Once user selects all the parameter required for prediction, a click action is required.
6. A single click by user will generate prediction result for a particular time span ranging from 15 days to 75 days .
7. Prediction range will be extended to a time period of advance 120 – 160 days.

Details about Some Vegetables

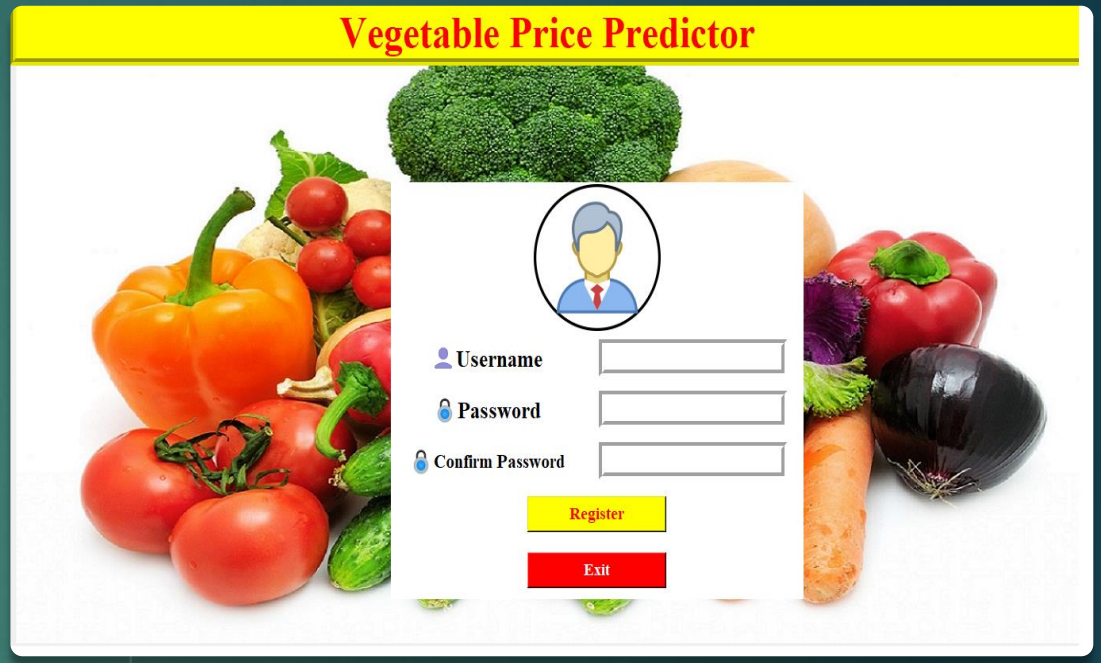
PARAMETERS	POTATO	TOMATO	ONION	GREEN CHILLY
Lifespan (Pantry)	1-2 Weeks	About 1 Week	4-6 Weeks	About 1 Week
Lifespan (Fridge/ Cold store)	7-11 Weeks	About 2 Week	4-8 Weeks	2-3 Weeks
Ready for Harvest in	8-12 Weeks*	8-10 Weeks*	12-20 Weeks#	4-5 Weeks
Season of Planting	January-February, October-November*	June-July, March-April	November-January, January -May	January – February, June – July, September- October
Soil	almost on any type of soil except saline and alkaline soils	wide range of soils from sandy to heavy clay	heavy soils, friable loam and alluvial soils	Well drained loamy soils rich in organic
Climatic Condition	20 – 24 ^{0} C	21-24°C.	13-24˚C 16-25˚C	

* varies from region to region/ climatic condition

depends upon usage as

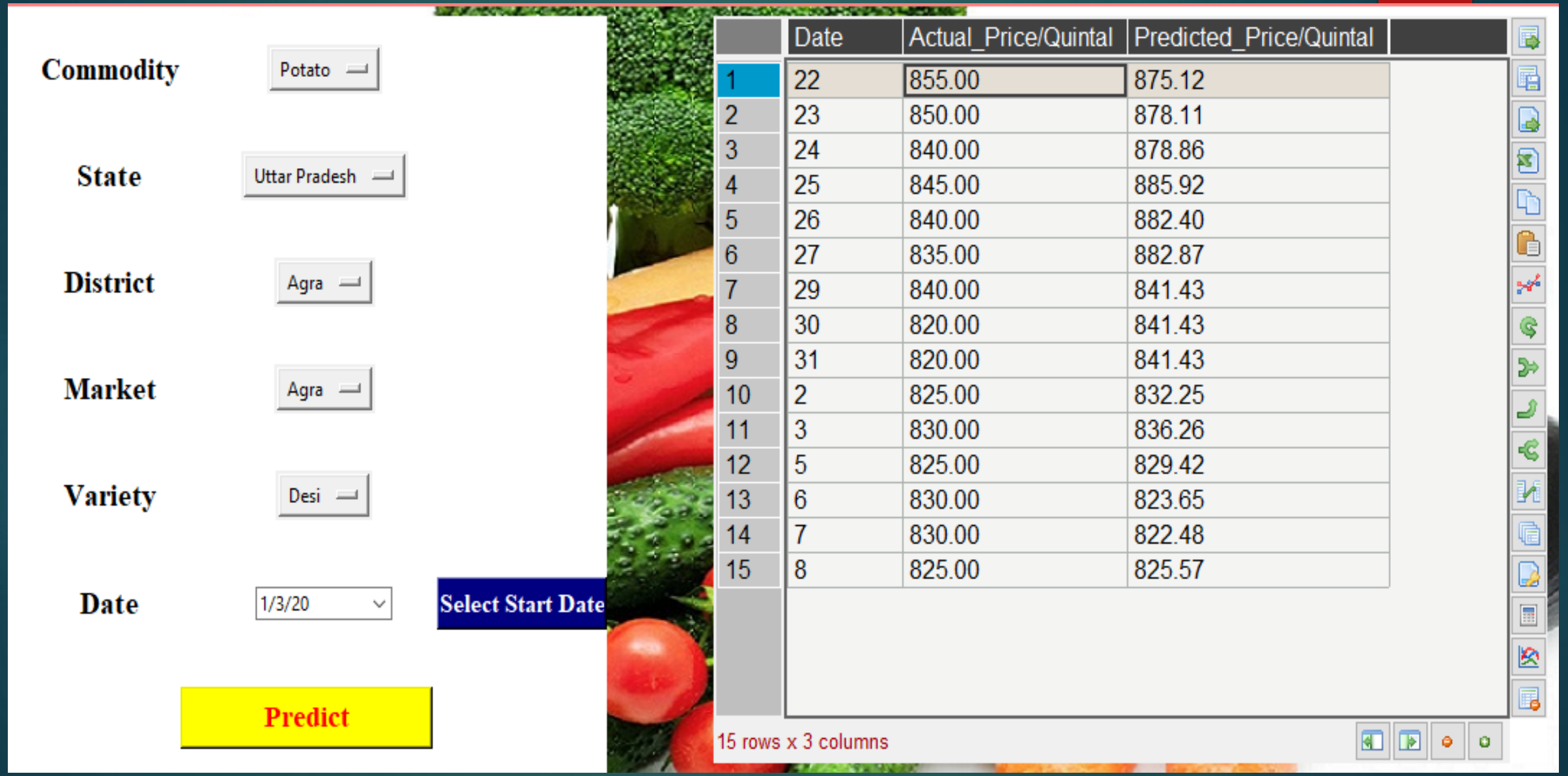
User Interface of System

Home Page



Signup Window

System in Action



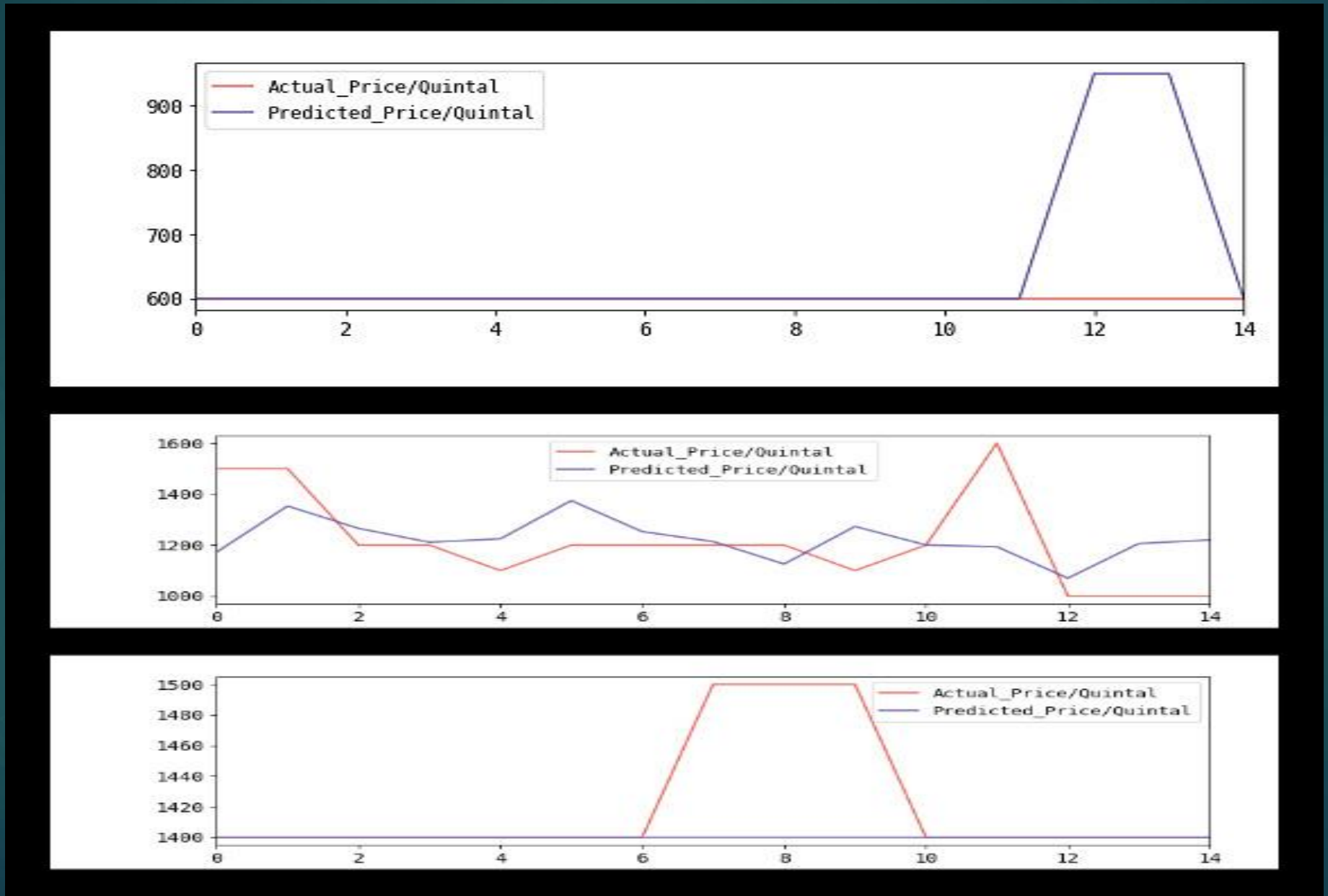
	Date	Actual_Price/Quintal	Predicted_Price/Quintal
1	22	855.00	875.12
2	23	850.00	878.11
3	24	840.00	878.86
4	25	845.00	885.92
5	26	840.00	882.40
6	27	835.00	882.87
7	29	840.00	841.43
8	30	820.00	841.43
9	31	820.00	841.43
10	2	825.00	832.25
11	3	830.00	836.26
12	5	825.00	829.42
13	6	830.00	823.65
14	7	830.00	822.48
15	8	825.00	825.57

15 rows x 3 columns

Result

State	District	Market	Accuracy
Odisha	Cuttack	Banki	1.0
Assam	Dibrugarh	Dibrugarh	0.99366
West Bengal	Howrah	Ramkrishnapur	0.95489
Madhya Pradesh	Dewas	Haatpipliya	0.93387
Uttar Pradesh	Agra	Agra	0.90819
Gujarat	Banaskanth	Dessa	0.84522
Himanchal Pradesh	Mandi	Mandi	0.75385
Maharashtra	Mumbai	Vashi New Mumbai	0.73408
Jharkhand	Giridih	Giridih	0.72347

Graph of Actual vs Predicted Price



Conclusion

Our system has been implemented using two datasets . One dataset contain parameters such as min_price, max_price, modal_price, arrival_date of vegetable . Other dataset contain weather parameters such as rainfall, humidity, temperature along with price & date. After applying algorithms on both datasets we come to know that dataset with more numbers of parameters produces higher accuracy as compared to dataset with lesser parameter. Average accuracy of the system is around 80-90%. For some place accuracy of the system reaches till 99%.

Framework & Library Used

- **Python 3.X**

- **Libraries Used**
  1. Pandas
  2. Numpy
  3. Skleran
  4. Tkinter

References

1. S. M. Idrees, M. A. Alam and P. Agarwal, ”A Prediction Approach for Stock Market Volatility Based on Time Series Data,” in IEEE Access, vol. 7, pp. 17287-17298, 2019.
2. G. M. Nasira and N. Hemageetha, ”Vegetable price prediction using data mining classification technique,” International Conference on Pattern Recognition, Informatics and Medical Engineering (PRIME-2012), Salem, Tamilnadu, 2012, pp. 99-102
3. Y. Lu, L. Yuping, L. Weihong, S. Qidao, L. Yanqun and Q. Xiaoli, ”Vegetable Price Prediction Based on PSO-BP Neural Network,” 2015 8th International Conference on Intelligent Computation Technology and Automation (ICICTA), Nanchang, 2015, pp. 1093-1096.
4. Luo C., Wei Q., Zhou L., Zhang J., Sun S. (2011) Prediction of Vegetable Price Based on Neural Network and Genetic Algorithm. In: Li D., Liu Y., Chen Y. (eds) Computer and Computing Technologies in Agriculture IV. CCTA 2010. IFIP Advances in Information and Communication Technology, vol 346. Springer, Berlin, Heidelberg

5. http://nhb.gov.in/PDFViwer.aspx?enc=3ZOO8K5CzcdC/Yq6HcdIxLoTxY3HgcNhvhmzPQYeGAU=
6. https://farmer.gov.in/
7. https://www.msamb.com/
8. https://food.ndtv.com/food-drinks/how-to-store-green-chillies-to-increase-their-shelf-life-1878544
9. http://www.dogr.res.in/index.php/en/onion/o-climent-and-soil
10. http://www.eatbydate.com/vegetables/fresh-vegetables/potatoes-shelf-life-expiration-date/

Thank You!