Crop price prediction in Indian markets using machine learning

Monika T
K Harsha Vamsi
Sowmya G
Indushekar A.C

Prof.Pooja Agarwal
Batch No. - 24

March 8, 2018

PES
Institute of Technology

Problem Statement / Definition

A tenacious problem that every farmer faces is not knowing when the ideal time to sell his crop would be. With advances being made everyday in the field of machine learning, it is only appropriate to use the ever growing technology to solve a common but rather underrated issue, predicting the future pricing of some of the most popular crops being grown in Indian states.
A mathematical model to predict potential global market crop prices can be incorporated into a user friendly user interface to give farmers the ability to be up to date on the current pricing.

Motivation Of The Work

A lot of farmers in India are unaware of the global market pricing and the fluctuations in them. This affects a lot of the decision making when it comes to buying and selling grains. We think that this can be changed by providing monthly updates on the current market scenario as well as graph visualizations to make it easier to analyze trends in the market.

Due to the growing digital revolution in India combined with the increasing Internet connectivity, this would be a high impact project with a wide reach across the country.

Literature Survey

- Yung-Hsing Peng, Chin-Shun Hsu, and Po-Chuang Huang in their research work on [1] Developing Crop Price Forecasting Service Using Open Data from Taiwan Markets aim to implement some well-known techniques for time series analysis, which are used to provide forecasting service for future crop prices.

  To track and to forecast the market prices are both important tasks in agri-management, by which the production schedule can be adjusted to increase the profit. For tracking the crop prices, the Council of Agriculture (COA) establishes an official website that provides open data of daily market prices from over 15 local markets with more than 100 different crops.

Literature Survey (contd)

- Manpreet Kaur, Hena Gulati and Harish Kundra in their research work on[2] Data Mining in Agriculture on Crop Price Prediction: Techniques and Applications mention some Data mining classification techniques can be used to develop an innovative model to predict the market price of respective commodity.

- Changshou Luo, Qingfeng Wei, Liying Zhou, Junfeng Zhang, and Sufen Sun in their research work on[3] Prediction of Vegetable Price Based on Neural Network and Genetic Algorithm explain the theory and construction methods of four models that are presented for predicting the vegetable market price, which are BP neural network model, the neural network model based on genetic algorithm, RBF neural network model and an integrated prediction model based on the three models above.

Methodology

- We first would need to formulate a list of Indian states and some of the most common crops grown in each state.
- Data curation would be the next step where we would use data available from the AgMarknet data hub to formulate our own dataset consisting of crop pricing data for the last 17 years. AgMarknet is an Indian government run initiative connecting farmers to the market through a digital platform.
- Once the dataset is built, the next step would be to build a mathematical machine learning model using the various input parameters like Max price, Min price, Avg price to produce a successful prediction of crop pricing on a monthly basis

Methodology (contd)

- This output can then be converted into a time series analysis problem and can be interactively shown to the farmer using graphs and charts.
- The end goal would be to build an Application which would be easily usable by the farmers to receive notifications on low crop prices and also keep them updated about the global market scenario for crops being grown by them.
- The last step would be to optimize this Application to automatically scale up to more locations and add more info as and when required.

Sample Data

State wise Wholesale Prices Monthly Analysis for Apple January, 2016

| State	Prices January, 2016	Prices December, 2015	Prices January, 2015	% Change(Over Previous Month)	% Change(Over Previous Year)
Assam	8227.27		3032	—	171.35
Bihar	6431.82	6416.67	3780	0.24	70.15
Gujarat	5498.74	5398.92	10309.89	1.85	-46.67
Haryana	5361.66	4236.49	5217.04	26.56	2.77
Himachal Pradesh	6023.16	5845.36	5800.67	3.04	3.84
Jammu and Kashmir	2230.14	2499.8	2463.48	-10.79	-9.47
Jharkhand	6393.42	6325.73	7405.74	1.07	-13.67
Karnataka	5961.06	6689.55	8419.93	-10.89	-29.2
Kerala	11333.56	9739.22	10640.9	16.37	6.51
Madhya Pradesh	2666.86	3500	2931.17	-23.8	-9.02
Maharashtra	3091.89	3507.87	3233.51	-11.86	-4.38
NCT of Delhi	1947.11	3041.64	2500	-35.98	-22.12
Orissa	8104.76	8364.32	8134.41	-3.1	-0.36
Punjab	3662.79	3235.12	4435.42	13.22	-17.42
Rajasthan	3878.31	4183.86	5444.86	-7.3	-28.77
Telangana	603.22	720.54	607.91	-16.28	-0.77
Tripura		1500		—	—
Uttar Pradesh	4911.33	4733.39	5385.3	3.76	-8.8
Uttrakhand	3450.67	3099.18	3896.25	11.34	-11.44
West Bengal	6438.42	5953.41	5804.2	8.15	10.93
Average	5064.01	4683.74	5233.83		

Workflow
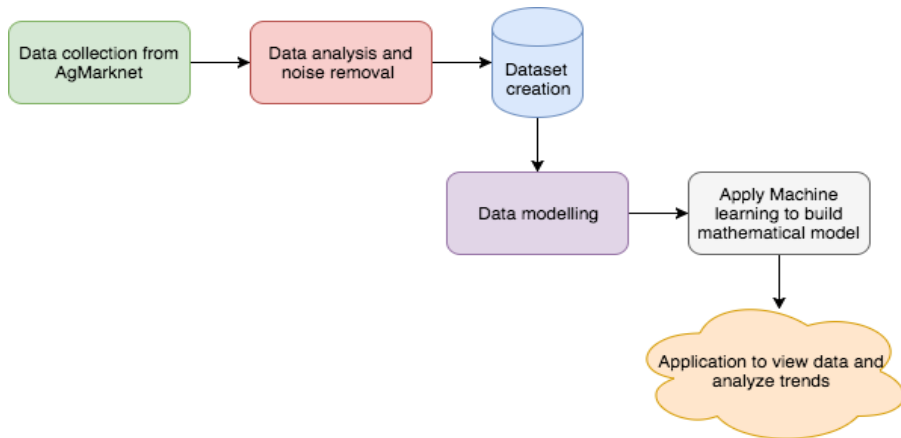
Actual Data

```
state,district,market,type,variety,arrivals,low,open,high,date
Andhra Pradesh,Krishna,Gudiwada,AP Raw Rice-1001,Cereals,100.00,1000,1100,1050,01 Jan 2008
Andhra Pradesh,Krishna,Gudiwada,B P T,Cereals,100.00,1750,1850,1800,01 Jan 2008
Andhra Pradesh,West Godavari,Kovvur,AP Raw Rice-1001,Cereals,0.01,1030,1050,1040,01 Jan 2008
Andhra Pradesh,West Godavari,Kovvur,Sona,Cereals,0.01,1680,1700,1690,01 Jan 2008
Andhra Pradesh,West Godavari,Tadepalligudem,Sona,Cereals,0.01,1700,1900,1800,01 Jan 2008
Assam,Kamrup,P.O. Uparhali Guwahati,Coarse,Cereals,110.00,1200,1500,1300,01 Jan 2008
Assam,Kamrup,P.O. Uparhali Guwahati,Fine,Cereals,110.00,1500,1900,1700,01 Jan 2008
Assam,Kamrup,P.O. Uparhali Guwahati,Fine(Basmati),Cereals,110.00,1200,2100,1700,01 Jan 2008
Bihar,Araria,Arreria,Medium,Cereals,5.00,1200,1800,1500,01 Jan 2008
Bihar,Bhabhua,Mohana,Medium,Cereals,50.00,950,1120,1035,01 Jan 2008
Gujarat,Anand,Khambhat(Grain Market),Fine,Cereals,0.20,1500,1600,1590,01 Jan 2008
Gujarat,Anand,Khambhat(Grain Market),Other,Cereals,0.20,1600,1600,1600,01 Jan 2008
Gujarat,Anand,Khambhat(Grain Market),Super Fine,Cereals,0.20,1700,1800,1785,01 Jan 2008
Gujarat,Dahod,Dahod,Other,Cereals,49.80,1500,1700,1600,01 Jan 2008
Gujarat,Panchmahals,Godhra,Medium,Cereals,32.60,1800,2000,1900,01 Jan 2008
Gujarat,Panchmahals,Godhra,Other,Cereals,32.60,1200,1400,1300,01 Jan 2008
Gujarat,Vadodara(Baroda),Vadodara,Kolam-Non Basmati,Cereals,143.20,1950,2000,2000,01 Jan 2008
Gujarat,Vadodara(Baroda),Vadodara,Masuri,Cereals,143.20,1350,1550,1450,01 Jan 2008
Gujarat,Vadodara(Baroda),Vadodara,Other,Cereals,143.20,1450,1650,1550,01 Jan 2008
Gujarat,Vadodara(Baroda),Vadodara,Parmal,Cereals,143.20,1250,1700,1500,01 Jan 2008
Jharkhand,Bokaro,Bokaro (Chas),Other,Cereals,14.00,1400,1540,1520,01 Jan 2008
Jharkhand,Deogarh,Madhupur,Coarse,Cereals,25.00,1100,1200,1200,01 Jan 2008
Jharkhand,East Singhbhum,Jamshedpur,Fine,Cereals,73.00,1550,1600,1550,01 Jan 2008
Jharkhand,East Singhbhum,Jamshedpur,Medium,Cereals,73.00,1400,1450,1400,01 Jan 2008
Jharkhand,Giridih,Giridih,Fine,Cereals,1.00,1600,1700,1700,01 Jan 2008
Jharkhand,Giridih,Giridih,Medium,Cereals,1.00,1200,1300,1300,01 Jan 2008
Jharkhand,Giridih,Giridih,Motta (Coarse) Boiled,Cereals,1.00,900,950,950,01 Jan 2008
Jharkhand,Gumla,Gumla,Other,Cereals,1.00,800,1200,1000,01 Jan 2008
Jharkhand,Koderma,Koderma,Fine,Cereals,18.00,1350,1500,1450,01 Jan 2008
Jharkhand,Koderma,Koderma,Medium,Cereals,18.00,1200,1400,1300,01 Jan 2008
Jharkhand,Pakur,Pakur,Coarse,Cereals,2.50,1150,1150,1100,01 Jan 2008
Jharkhand,Pakur,Pakur,Fine,Cereals,2.50,1650,1750,1700,01 Jan 2008
Jharkhand,Pakur,Pakur,Medium,Cereals,2.50,1350,1450,1400,01 Jan 2008
Jharkhand,Ranchi,Ranchi,Coarse,Cereals,107.50,1115,1160,1140,01 Jan 2008
Jharkhand,Ranchi,Ranchi,Fine,Cereals,107.50,1920,1975,1925,01 Jan 2008
Jharkhand,Ranchi,Ranchi,Medium,Cereals,107.50,1340,1370,1350,01 Jan 2008
Jharkhand,Saraikela(Kharsawan),Sariakela,Medium,Cereals,25.00,1125,1175,1125,01 Jan 2008
Jharkhand,West Singbhum,Chaibasa,Coarse,Cereals,8.50,950,1050,1000,01 Jan 2008
Jharkhand,West Singbhum,Chaibasa,Fine,Cereals,8.50,1250,1350,1300,01 Jan 2008
Jharkhand,West Singbhum,Chaibasa,Medium,Cereals,8.50,1150,1250,1200,01 Jan 2008
Karnataka,Kolar,Bangarpet,Broken Rice,Cereals,183.00,600,720,700,01 Jan 2008
Karnataka,Kolar,Bangarpet,Medium,Cereals,183.00,1600,2000,1800,01 Jan 2008
Kerala,Malappuram,Thirurrangadi,Boiled Rice,Cereals,12.00,1800,1900,1850,01 Jan 2008
Kerala,Malappuram,Thirurrangadi,Other,Cereals,12.00,1750,1850,1800,01 Jan 2008
Kerala,Malappuram,Thirurrangadi,Pusa Basmati Raw (Old),Cereals,12.00,3700,4300,4000,01 Jan 2008
```
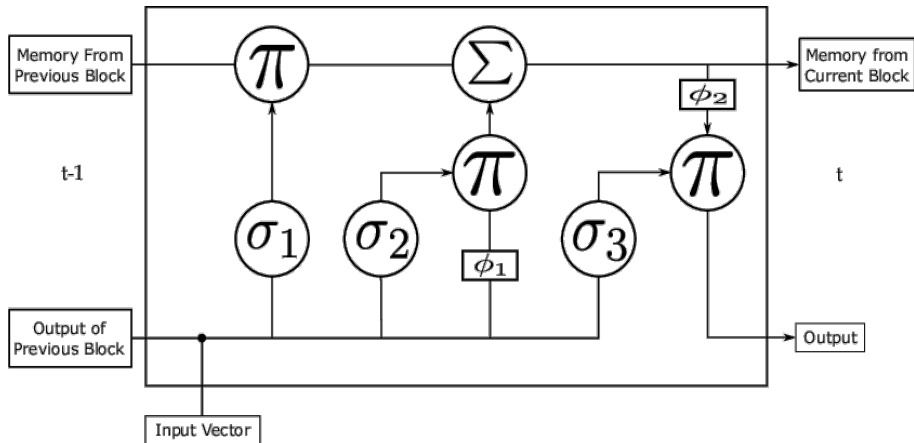
Data Visualization
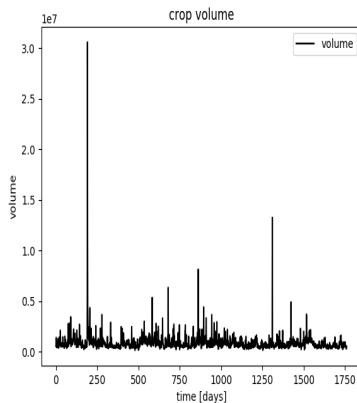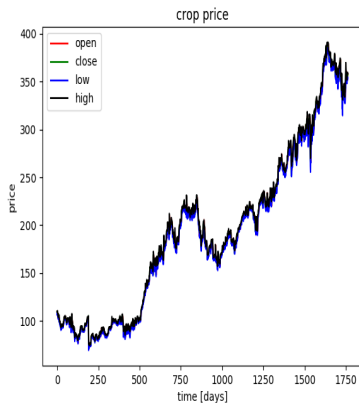
Normalized Data Visualization
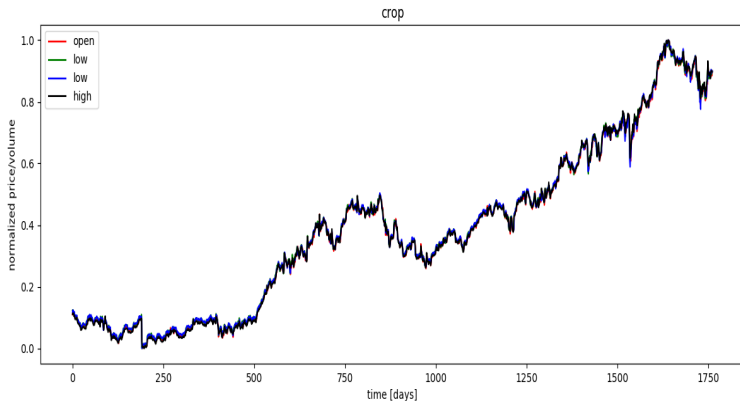
Detailed Design of LSTM

# LSTM Model With and Without Sequence (Methodology)

- We train an LSTM model on this data with and without sequence.
- We make use of the closing price for training and validation. We use the relu activation function with 5 epochs.
- The look back period is 7 days and the look forward period is 1 day.
- The loss represents the mean squared error and is 0.0022 for sequence data and 0.0070 for sequence-less data.

```
Epoch 1/5                                   Epoch 1/5
3s - loss: 0.0326 - val_loss: 0.0067        2s - loss: 0.0391 - val_loss: 0.0079
Epoch 2/5                                   Epoch 2/5
1s - loss: 0.0033 - val_loss: 7.8516e-04    1s - loss: 0.0162 - val_loss: 0.0026
Epoch 3/5                                   Epoch 3/5
1s - loss: 0.0023 - val_loss: 4.7136e-04    1s - loss: 0.0103 - val_loss: 0.0020
Epoch 4/5                                   Epoch 4/5
1s - loss: 0.0025 - val_loss: 3.0357e-04    1s - loss: 0.0097 - val_loss: 0.0097
Epoch 5/5                                   Epoch 5/5
1s - loss: 0.0022 - val_loss: 2.6147e-04    1s - loss: 0.0070 - val_loss: 0.0012
```

Output Loss Visualization

Comparison of log loss for sequence and sequence-less data
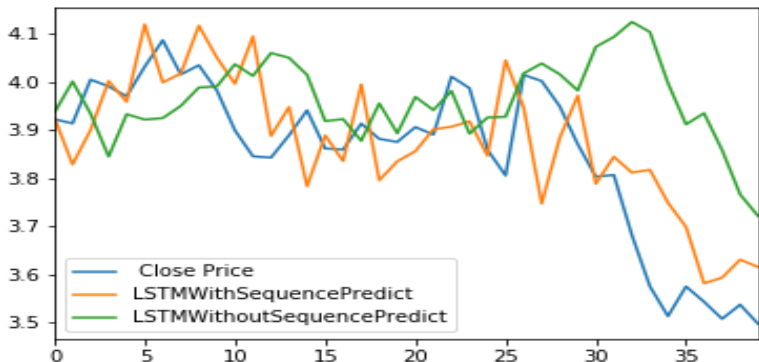
Time-line of Completion of Project from Nov 2017-10th April 2018(Gantt Charts).
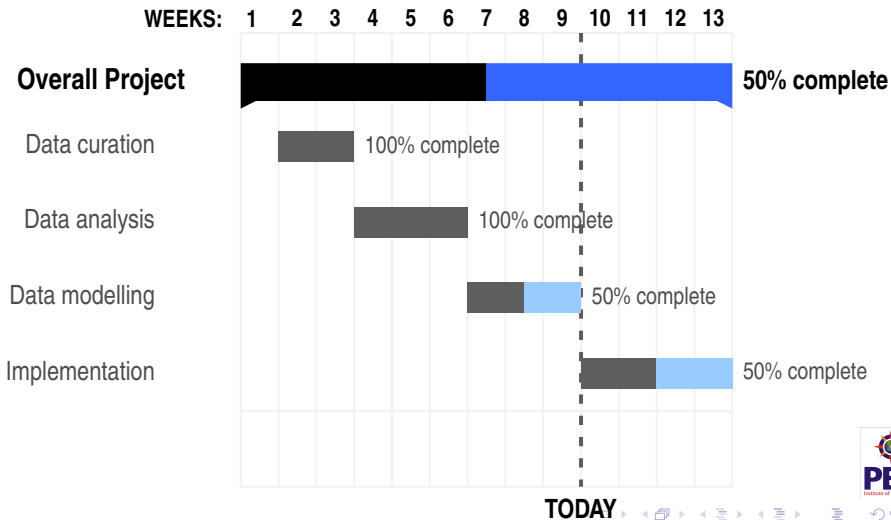
Expected Outcome/ Results

1.Estimating crop prices

Our end goal would be to accurately estimate prices for a chosen crop for any given month and hence help the farmer in making wise decisions while selling/buying from the wholesale market.

2.A simple User Interface

We aim to develop an Application to reach a wider farmer crowd due to growing internet connectivity in most of the Indian states. This would give farmers a quick way to access data without needing to visit the market frequently.

3.Optimizing on device experience

We would also provide graphs and charts to show time-series data over a period of a year to give an idea of the crop pricing trends.

References

📄 Yung-Hsing Peng, Chin-Shun Hsu, and Po-Chuang Huang
Developing Crop Price Forecasting Service Using Open Data from Taiwan Markets
Institute for Information Industry, Kaohsiung City, Taiwan.

📄 Manpreet Kaur, Heena Gulati and Harish Kundra
Data Mining in Agriculture on Crop Price Prediction: Techniques and Applications
International Journal of Computer Applications (0975 8887) Volume 99 No.12.

📄 Changshou Luo, Qingfeng Wei, Liying Zhou, Junfeng Zhang, and Sufen Sun
Prediction of Vegetable Price Based on Neural Network and Genetic Algorithm
Institute of Information on Science and Technology of Agriculture, Beijing Academy of Agriculture and Forestry Sciences, Beijing.

The End