
Analyze Crop Production Of India
SPOTLE.AI -NASSCOM Community
AI Internship Program
By Anindya Nag

Declaration

We certify that

- The work contained in this report is original and has been done by me.
- I have followed the guidelines provided by the authority in preparing the report.
- Whenever, I have used materials (data, theoretical analysis, figures, and text) from other sources, I have given due credit to them by giving their details in the references.
- Neither this project nor any part of it has been submitted for any degree or academic award elsewhere.

Anindya Nag

Signature

Acknowledgement

First and foremost, I would like to express deepest gratitude to the authority spotle.ai – NASSCOM Community to give us such a good project.

We are thankful to all the faculty members of the Department of CSE for their helpful comments.

Last but not the least, we would like to thank our friends for their encouragement and help in several forms.

Anindya Nag

Date: 14.07.2021

Abstract

Ensuring food security ought to be an issue of great importance for a country like India where more than one-third of the population is estimated to be absolutely poor and one-half of all children are malnourished in one way or another. Within season crop production forecasts are widely recognized as an important input in analyzing food balance sheets and anticipating production shortfalls. Though crop production estimation and assessment are done worldwide on a regional extent, advance yield prediction over space and lead-times is less popular especially in India. Limited spread of observatories, lack of infrastructure in the observatories, dynamicity of weather, availability of less efficient process-based approaches to predict the turbulence of weather, heterogeneity in agriculture, lacking in integration of processes, etc. pose constraints making it a risky field with not much effective methodology developed till date. In past, the most attempts are made to forecast the crop yield in purely statistical and semi-statistical basis, which proved to be very biased to the location and the year they are developed.

This project aims at developing a more scientific approach to forecast the crop yield on regional scale and at various time leads within the growing season. The objective is to develop a robust methodology to forecast the yield at high resolution spatially and temporally with a known level of accuracy and build a prediction model of crop production.

Contents

Introduction

Agriculture produce is subjected to various risks, which are not only confined to production risk pertaining to weather, pest but also the demand and supply of various countries, other policy and economic factors. With restricted knowledge to understand and comprehend the information, farmers can incur huge losses by selling their produce in distress. Farmers no longer have to contend with just local markets. They also have to account for competition from the world over.

The authority gave us a dataset contains information on crop covered area (Hectare) and production (Tonnes) for 122 different crops in 33 states of India across 14 years (2000-2013). Using this dataset, we have to study and analyses crop production, production contribution to State/country, performance, and high yield production order for crops, crop growing pattern and diversification. Also, you have to forecast the product of the crop for future periods, which can be used to formulate crop-related schemes.

Pre-requisites

Prerequisites for our project are –

a) python3 or jupyter notebook should be installed.

b) Following packages should be installed –

- cv2
- NumPy
- pandas
- matplotlib. pyplot
- tflearn

c) **Crop Prediction Dataset**: *https://cdn.spotle.ai/datasets/state-wise-crop-production-India.zip*

Steps for the Code

- **Build a model to predict the production of Rice in 2014 for Andhra Pradesh, Uttar Pradesh and Tamil Nadu.**

Step–1: Importing Libraries

In the first step, we will be importing all the respective libraries as mentioned above.

```
In [1]: import numpy as np
        import pandas as pd
        import matplotlib.pyplot as plt
        %matplotlib inline
        import seaborn as sns
        #to ignore the warnings
        import warnings as wg
        wg.filterwarnings("ignore")
```

Step-2: Importing the Dataset

In this step, we will import our given dataset.

```
In [3]: #reading data from file
        url="https://cdn.spotle.ai/datasets/state-wise-crop-production-India.zip"
        df=pd.read_csv(url)
```

```
In [4]: df
```

Out[4]:

	State	Year	Crop	Area	Production

Step-3: Creating Dataframe

According to the above question, we have to build a model for predicting rice production in 2014 for Andra Pradesh, Uttar Pradesh, Tamil Nadu. For this, we have to create a dataframe for rice production.

```
:  rice_df=df[df['Crop']=='Rice']
```

Now we have to create a dataframe for Andra Pradesh Rice Production.

```
In [29]: rice_df=df[df['Crop']=='Rice']
         andhra_pradesh_rice_production_df=rice_df[df['State']=='Andhra Pradesh']
         andhra_pradesh_rice_production_df
```

Out[29]:

	State	Year	Crop	Area	Production
121	Andhra Pradesh	2000	Rice	2694741.0	8040667.0
162	Andhra Pradesh	2001	Rice	2515353.0	7823692.0
220	Andhra Pradesh	2002	Rice	1867142.0	5315408.0
277	Andhra Pradesh	2003	Rice	1957323.0	6054099.0
318	Andhra Pradesh	2004	Rice	2227682.0	7392678.0
357	Andhra Pradesh	2005	Rice	2520723.0	7288721.0
396	Andhra Pradesh	2006	Rice	2488883.0	7616392.0
425	Andhra Pradesh	2007	Rice	2576000.0	8880000.0
458	Andhra Pradesh	2008	Rice	2695725.0	8880586.0
496	Andhra Pradesh	2009	Rice	2325729.0	7569071.0
536	Andhra Pradesh	2010	Rice	2772377.0	7883078.0
578	Andhra Pradesh	2011	Rice	2346300.0	7744175.0
624	Andhra Pradesh	2012	Rice	2209237.0	6862854.0
673	Andhra Pradesh	2013	Rice	2583440.0	7993425.0

Step-4: Code for Uttar Pradesh & Tamil Nadu:

Similar way we have to do the same thing for Uttar Pradesh & Tamil Nadu.

```
In [27]: rice_df=df[df['Crop']=='Rice']
uttar_pradesh_rice_production_df=rice_df[df['State']=='Uttar Pradesh']
uttar_pradesh_rice_production_df
```

Out[27]:

	State	Year	Crop	Area	Production
9365	Uttar Pradesh	2000	Rice	5907151.0	11679149.0
9388	Uttar Pradesh	2001	Rice	6071325.0	12855857.0
9414	Uttar Pradesh	2002	Rice	5213429.0	9596346.0
9441	Uttar Pradesh	2003	Rice	5727772.0	12494939.0
9466	Uttar Pradesh	2004	Rice	5934405.0	10749989.0
9492	Uttar Pradesh	2005	Rice	5571524.0	11119095.0
9524	Uttar Pradesh	2006	Rice	5920752.0	11123613.0
9554	Uttar Pradesh	2007	Rice	5709388.0	11780112.0
9586	Uttar Pradesh	2008	Rice	6033382.0	13097031.0
9618	Uttar Pradesh	2009	Rice	5172741.0	10776504.0
9648	Uttar Pradesh	2010	Rice	5657070.0	11992299.0
9681	Uttar Pradesh	2011	Rice	5946285.0	14022392.0
9714	Uttar Pradesh	2012	Rice	5861282.0	14415939.0
9747	Uttar Pradesh	2013	Rice	5982369.0	14635836.0

```
In [28]: rice_df=df[df['Crop']=='Rice']
tamil_nadu_rice_production_df=rice_df[df['State']=='Tamil Nadu']
tamil_nadu_rice_production_df
```

Out[28]:

	State	Year	Crop	Area	Production
8133	Tamil Nadu	2000	Rice	2113344.0	7217903.0
8151	Tamil Nadu	2001	Rice	2059878.0	6583630.0
8214	Tamil Nadu	2002	Rice	1516537.0	3577108.0
8292	Tamil Nadu	2003	Rice	1396651.0	3222776.0
8331	Tamil Nadu	2004	Rice	1872822.0	5061622.0
8368	Tamil Nadu	2005	Rice	2050455.0	5209433.0
8404	Tamil Nadu	2006	Rice	1931397.0	6610607.0
8439	Tamil Nadu	2007	Rice	1789170.0	5039954.0
8473	Tamil Nadu	2008	Rice	1931603.0	5183385.0
8507	Tamil Nadu	2009	Rice	1845553.0	5665258.0
8526	Tamil Nadu	2010	Rice	1905726.0	5792415.0
8555	Tamil Nadu	2011	Rice	1903772.0	7458657.0
8574	Tamil Nadu	2012	Rice	1493276.0	4050334.0
8602	Tamil Nadu	2013	Rice	1725730.0	7115195.0

Step-5: Splitting the data into train & test:

Now, we will split our dataset into train data & test data.

```
In [32]: crop_data=pd.get_dummies(data=df)
```

```
In [33]: from sklearn.model_selection import train_test_split
```

```
In [34]: X=crop_data.drop('Production',axis=1)
X.head()
```

Out[34]:

	Year	Area	State_Andaman and Nicobar Islands	State_Andhra Pradesh	State_Arunachal Pradesh	State_Assam	State_Bihar	State_Chandigarh	State_Chhattisgarh	State_Dadra and Nagar Haveli	...	Crop_Turm
0	2000	4354.0	1	0	0	0	0	0	0	0	...	
1	2000	1707.0	1	0	0	0	0	0	0	0	...	
2	2000	800.0	1	0	0	0	0	0	0	0	...	
3	2000	25160.0	1	0	0	0	0	0	0	0	...	
4	2000	388.0	1	0	0	0	0	0	0	0	...	

5 rows × 157 columns

```
In [35]: y=crop_data['Production']
         y.head()

Out[35]: 0        7200.0
         1       12714.0
         2         219.0
         3    89000000.0
         4        1220.0
         Name: Production, dtype: float64
```

Step- 6: Training the dataset:

We have spilted our dataset and now, we will train our model.

```
from sklearn.linear_model import LinearRegression
crop_model=LinearRegression()
crop_model.fit(X_train,y_train)

Out[38]: LinearRegression()
```

```
In [40]: print(X_test)
         y_pred=crop_model.predict(X_test)
         y_test

              Year      Area  State_Andaman and Nicobar Islands  \
         2882  2008   175400.0                                  0
         3933  2000  1022701.0                                  0
         304   2004  1646166.0                                  0
         3325  2011    41197.0                                  0
         2626  2010     7370.0                                  0
         ...    ...       ...                                ...
         2102  2007       50.0                                  0
         997   2005    17875.0                                  0
         9648  2010  5657070.0                                  0
         9891  2004       21.0                                  0
         1176  2011    65560.0                                  0

              State_Andhra Pradesh  State_Arunachal Pradesh  State_Assam  State_Bihar  \
         2882                    0                        0            0            0
         3933                    0                        0            0            0
         304                     1                        0            0            0
         3325                    0                        0            0            0
         2626                    0                        0            0            0
         ...                   ...                      ...          ...          ...
         2102                    0                        0            0            0
         997                     0                        0            1            0
         9648                    0                        0            0            0
         9891                    0                        0            0            0
```

```
      Crop_other oilseeds
2882                    0
3933                    0
304                     0
3325                    0
2626                    0
...                   ...
2102                    0
997                     0
9648                    0
9891                    0
1176                    0

[3533 rows x 157 columns]
```

Out[40]:
```
2882     176800.0
3933    1835332.0
304     1442146.0
3325     149000.0
2626       7836.0
           ...
2102         68.0
997       11309.0
9648   11992299.0
9891          7.0
1176     608023.0
Name: Production, Length: 3533, dtype: float64
```

In [41]:
```
y_pred
```

Out[41]:
```
array([  2753862.88178504, -13146110.92409635,  23793940.23387146, ...,
        63505173.89426923,   7680127.66949213, -11892716.04433429])
```

In [42]:
```
#comparing acrual vs predicted
comp=pd.DataFrame({'Actual':[y_test],'Predicted':[y_pred]})
comp
```

Out[42]:

	Actual	Predicted
0	2882 176800.0 3933 1835332.0 304 ...	[2753862.881785035, -13146110.924096346, 23793...

4.Report

Introduction:

28 Rice is grown in almost all the districts of Andhra Pradesh. However, the major rice producing districts of Andhra Pradesh are situated in the fertile Krishna Godavari basin region. The districts that have earned the name "the rice bowl of India" are West Godavari, East Godavari and Krishna. Besides the in the irrigated regions of Telangana, namely in the district of Karimnagar too, rice is cultivated as a major crop. Despite the predominance of agriculture as a major source of livelihood, there are certain problems facing the farmers which need state intervention, specifically in the problems related to flash floods or droughts, deteriorating soil health, lack of access to formal credit sources, maintenance of irrigation structures and extension services.

Uttar Pradesh has 5.6 million ha under rice cultivation, which covers irrigated and rainfed areas. The area has been fluctuating between 5.2 and 6.1 million ha during the last seven years. The current state average productivity is about 2 tonnes/ha and the major constraints in production are low coverage by hybrids and varieties, soil salinity/alkalinity and micro and secondary nutrient deficiency, imbalanced fertilizer use, decreasing organic carbon content in soils; flash floods, submergence and drought in eastern part and borers.

Tamil Nadu has 2.2 million ha under rice cultivation, which covers mainly irrigated and partly rainfed areas. The state average productivity is about 2.8 tones/ha. The major constraints in production are water scarcity in the existing command areas, imbalanced fertilizer use and micronutrient deficiency and increasing soil salinity/alkalinity.

Growth of Paddy Crop in Andhra Pradesh:

The state has about 4 million ha under rice which is mainly irrigated. The area has been fluctuating between 2.8 and 4.2 million ha during the last seven years. The state average productivity is about 3 tonnes/ha. The major constraints in production are biotic stresses such as bacterial leaf blight (BLB), gall midge and sheath blight and abiotic stresses like water scarcity.

Growth of Paddy Crop in Uttar Pradesh:

In most of UP. there is adequate surface water during the rainy season and shallow depths of ground water. These imply one or two supplementary irrigations as feasible and economical, but seldom practiced Somehow, the tendency is to wait for the rains. The farmers need to be educated to change this attitude. For example, there is no reason for drought in north eastern U.P. The water table is stable and just 3-5 meters below. Furthermore, during the rainy season there is adequate water in ponds and lakes for up to two irrigations to the entire rice areas planted Still almost every year the crop suffers from drought.

The government should encourage small irrigation projects and restoration of old ponds etc.: In certain areas use of mantul pumps possibly in combination with small tanks, Azolla and fish appears feasible. This could be an interim answer till some other source of power becomes available. Use of alternate sources of energy e.g., biogas, gasifiers, wind mills, photovoltaic cells (incider tally India is the largest producer of photovoltaic cells in the world) etc. should be explored.

Growth of Paddy Crop in Tamil Nadu:

Details of the foodgrains production in Tamil Nadu during1980-81 to 2012-13 are given in Table 3. The area underfoodgrain cultivation increased from 5487 million ha. in1980-81 to 8617 m. ha. in 2000-01. Thereafter, it declined to 6294 m. ha in 2012-13. But during 2011-12, the area under foodgrains was 9640.6 ha. The area under paddy crop occupies highest share among foodgrains in Tamil Nadu. The area under paddy crop increased from 75.80 percent(4159 m. ha.) in 1980-81 to 85.49 percent (7366 m. ha.) in2000-01. Thereafter, it is declined to 69.90 percent (4399m. ha.) in 2012-13. The area under paddy is the highest due to the adoption of modern technology in agriculture.

5.Conclusion

Crop yield prediction is still remaining as a challenging issue for farmers. The aim of this research is to propose and implement a rule-based system to predict the crop yield production from the collection of past data. This has been achieved by applying association rule mining on agriculture data from 2000 to 2013.

6.References

https://www.researchgate.net/publication/273691933_Rice_Production_Productivity_in_Andhra_Pradesh

https://en.wikipedia.org/wiki/Rice_production_in_India

https://books.google.co.in/books/about/Rice_Production_in_Uttar_Pradesh.html?id=Xx0XOtRVABoC&redir_esc=y

https://www.researchgate.net/publication/343306461_Growth_of_Rice_Production_in_Tamil_Nadu_Progress_and_Prospects