An AI Based Web Portal for Cotton Price Analysis and Prediction

Mohammed Bohara¹,*, Krishna Patel², Bansari Patel³, Jesal Desai⁴

¹Department of Computer Engineering, Devang Patel Institute of Advance Technology and Research, Charotar University of Science and Technology (CHARUSAT), CHARUSAT campus, Changa 388421, India.
²Department of Computer Science & Engineering, Devang Patel Institute of Advance Technology and Research, Charotar University of Science and Technology (CHARUSAT), CHARUSAT campus, Changa 388421, India.
³Department of Computer Science & Engineering, Devang Patel Institute of Advance Technology and Research, Charotar University of Science and Technology (CHARUSAT), CHARUSAT campus, Changa 388421, India.
⁴Department of Computer Science & Engineering, Devang Patel Institute of Advance Technology and Research, Charotar University of Science and Technology (CHARUSAT), CHARUSAT campus, Changa 388421, India.
*Corresponding author. Email: mohammedbohara.ce@charusat.ac.in

ABSTRACT

Artificial Intelligence based Web portal is extensively used in worldwide because it used to solve various primary low level problems in agriculture such as predicting precise prices of agriculture product using machine learning algorithms, enabling direct connection between buyer and seller without any middlemen, decreasing wastage of post-harvest crop, fixing appropriate price margin for farmers and so on. These steps helps to increase the capital income and creates the beneficial changes in the agriculture and also it leads to take the agriculture to next level. Cotton is a crop that is possibly the most expensive including all agricultural products sold. Owing to the high uncertainty of cotton markets, it is incredibly difficult to forecast future demand patterns and develop an acceptable pricing plan in order to compete in the market and maximize the company's volume for long-term development. Keeping in mind all the factors stated above for the cultivation of cotton, conclusion can be inferred as the cotton is highly volatile in nature. In this paper, we tend to build an Artificial Intelligence based Web Portal which provides data analysis of price of past years and also gives predictions about the cotton prices to the fellow traders. The paper aims to create a web model for cotton price prediction. Furthermore, the study tries to run a multiple regression model based on different variables to estimate the price of cotton.

Keywords: Artificial Intelligence, Agriculture, ARIMA Model, Cotton Price Analysis, Integration, Multiple Regression, SARIMA.

1. INTRODUCTION

India is the world's largest cotton producer and the world's second-largest cotton exporter. India produces around a quarter of all cotton in the country. Cotton is a global commodity with high market uncertainty that is affected by global market cycles. It's mostly used as a raw material in the clothing and textile industries. Cotton farmers face a high price risk in addition to output risk. As a result, it's vital to predict cotton prices for the good of both farmers and millers who buy cotton [1-4]. The aim of this research is to forecast cotton prices in India's major producing states. Cotton production accounts for approximately 9% of India's total agricultural property. Cotton crops account for 14-16 percent of India's overall agri-crop.

Cotton production is directly responsible for the livelihood of 60 million individuals, including 4.5 million farmers [5]. Keeping in mind all the factors stated above for the cultivation of cotton, conclusion can be inferred as the cotton is highly volatile in nature. In this paper, we intend to create an Artificial Intelligence- driven approach that can forecast future demand patterns based on historical data and propose a suitable pricing model for cotton sales [6-8].

As we know that, India is an Agro-centric country and more than half of India’s economy depends on...
agriculture. And almost 1300 million people are directly or indirectly dependent on the farming sector. Cotton is a major Agri-crop, as in production as well as consumption [9]. As a result, it has a major effect on the Indian agriculture sector as a whole and so, to bring a solution to the issues coming out of the farming sector becomes a necessity and because of the continuous fluctuation in the prices of cotton and for cotton-related businesses, staying on top of industry dynamics is critical. Until pursuing his estimate by agreeing to purchase or sell, traders must combine bits of knowledge about supply and demand and compare the guesses of fellow traders, i.e., what are the latest market trends? What are the forecasts for the future? Proposed approach for this problem is Artificial Intelligence based solution so as to analyse the past year’s data of the cotton crop cultivation (respective to the region – a particular market) and analysing the current trends and developing a model that will help the fellow market traders to predict the future market trends i.e., to suggest a suitable pricing Model [10-15].

The aim of this research is to construct a web model based on economic variables that can predict the future cotton price level more accurately than a random walk model and an autoregressive model [16]. Although forecasting is used in a broad range of different fields, much of the subject’s economic research is focused on forecasting inflation and exchange rates, whereas forecasting is based on the commodity market [17-19]. In three ways, cotton forecasting is especially important. First of all, it is a commodity that is widely used globally in industries that have a major impact on the economy of primarily third-world countries, where predictability of prices is important for stable and long-term growth. Secondly, since such contracts are of great benefit to producers, consumers, exporters and importers of goods, they are useful for the pricing and understanding of hedging instruments [20].

2. RELATED WORKS

The past studies of cotton price prediction are based on some of the cotton futures. Since, the farmers are working hard to earn the high profit from the crop. But, the farmers are failed to achieve profit, due to the transportation cost, traders and intermediator [21]. Therefore, the recommendation for market rate, price prediction and significant insights are generated by using the Big Data Analytics, Data Mining and Machine Learning [22]. In addition, it is proven that short-term futures are better [23,24]. Multiple empirical studies It asserts that the cost of cotton is found in the price of cotton futures. The price forecast at maturity relative to long-term futures [25] and sudden Changes in futures prices are expected to make an impact on the spot price for approximately about five months or more. Majority of with a one-month lag. Ashby and Brosen et al. suggested that a large amount of the relation between the price of futures and the spot prices of commodities. Multiple empirical studies it asserts that the cost of cotton is found in the price of cotton futures. In addition, it is proven that short-term futures are better [26]. The price forecast at maturity relative to long-term futures and sudden Changes in futures prices are expected to make an impact on the spot price for approximately about five months or more. Majority of with a one-month lag [27]. Full, however, forecast the future price of cotton, dependence on cotton futures can be conducive to significant failures of it [28]. In the opinion of Bernake (2008), quotes from upcoming markets were underestimated when this can contribute to subsequent under-predictions of overall inflation as a result of the pace of commodity price gains [29]. And in this regard, a new model is needed in order to precisely predict the cotton prices exploit, aside from supply and demand. The Price of Futures (ibid). In addition, previous studies have attempted to predict the price of cotton based on the quality of it. In addition, all macroeconomic variables, for an instance interest rates, exchange rates etc. have an effect on the macroeconomic situation. The price is omitted and varies over time. In the analysis, it is indicated that in a particular given the quality of the year is considered to adjust the price, the level of which depends on the price [30]. The characteristics of quality of the products available in the crop of that year and the demand for the features [31].

The attributes present in each area in the international cotton market. It can vary and the demand can adapt according to regions, so variables need to be modified. To be tailored to each area and each period of time. In addition, a study of Cho's Expands the field of analysis and adds to the model weather variables, enhancing the cotton Prediction of spot prices, based on some supply factors such as (quality and weather) [32-35]. The two-research hedonic regression model applied by Cho and Ethridge and Davis[36,37], an Alternate type of multiple regression model of ordinary least squares (OLS) to estimate Based on categorical variables, the price of cotton While both of the two tests are related to attributes of cotton quality [38-42].

3. PROPOSED MODELS

Proposed approach to solve this problem using Artificial Intelligence based solution so as to analyse the past year’s data of the cotton crop cultivation (respective to the region – a particular market) and analysing the current trends and developing a model that will help the fellow market traders to predict the future market trends i.e., to suggest a suitable pricing Model.

3.1. Dataset

A scraper script will successfully be scraped past 11 years of data regarding prices of cotton of each market in Gujarat. There are nearly 400 markets in Gujarat out of
which 77 Markets are Cotton Traders and Sellers. Here is an image of data of balasinor Market data [19,20,21].

Table 1. Dataset of Balasinor Market, Gujarat

<table>
<thead>
<tr>
<th>si</th>
<th>district</th>
<th>market</th>
<th>commodity</th>
<th>variety</th>
<th>min_p</th>
<th>max_p</th>
<th>min_r</th>
<th>max_r</th>
<th>mode</th>
<th>price</th>
<th>price</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Kheda</td>
<td>Balasinor</td>
<td>Cotton</td>
<td>RCH-2</td>
<td>4300</td>
<td>4500</td>
<td>4500</td>
<td>1/13/2020</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Kheda</td>
<td>Balasinor</td>
<td>Cotton</td>
<td>RCH-2</td>
<td>4000</td>
<td>4500</td>
<td>4500</td>
<td>1/16/2020</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Kheda</td>
<td>Balasinor</td>
<td>Cotton</td>
<td>RCH-2</td>
<td>4400</td>
<td>4600</td>
<td>4850</td>
<td>1/18/2020</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Kheda</td>
<td>Balasinor</td>
<td>Cotton</td>
<td>RCH-2</td>
<td>4200</td>
<td>4700</td>
<td>4700</td>
<td>1/23/2020</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Kheda</td>
<td>Balasinor</td>
<td>Cotton</td>
<td>RCH-2</td>
<td>4300</td>
<td>4600</td>
<td>4600</td>
<td>1/24/2020</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Kheda</td>
<td>Balasinor</td>
<td>Cotton</td>
<td>RCH-2</td>
<td>4300</td>
<td>4600</td>
<td>4600</td>
<td>1/27/2020</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Kheda</td>
<td>Balasinor</td>
<td>Cotton</td>
<td>RCH-2</td>
<td>4500</td>
<td>4600</td>
<td>4600</td>
<td>1/29/2020</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Kheda</td>
<td>Balasinor</td>
<td>Cotton</td>
<td>RCH-2</td>
<td>4000</td>
<td>5150</td>
<td>5150</td>
<td>12/29/2018</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Kheda</td>
<td>Balasinor</td>
<td>Cotton</td>
<td>RCH-2</td>
<td>4500</td>
<td>5150</td>
<td>5150</td>
<td>12/28/2018</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Kheda</td>
<td>Balasinor</td>
<td>Cotton</td>
<td>RCH-2</td>
<td>4500</td>
<td>5150</td>
<td>5150</td>
<td>12/27/2018</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Kheda</td>
<td>Balasinor</td>
<td>Cotton</td>
<td>RCH-2</td>
<td>4600</td>
<td>5150</td>
<td>5150</td>
<td>12/26/2018</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Kheda</td>
<td>Balasinor</td>
<td>Cotton</td>
<td>RCH-2</td>
<td>4600</td>
<td>5150</td>
<td>5150</td>
<td>12/25/2018</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Kheda</td>
<td>Balasinor</td>
<td>Cotton</td>
<td>RCH-2</td>
<td>4500</td>
<td>5150</td>
<td>5150</td>
<td>12/24/2018</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Kheda</td>
<td>Balasinor</td>
<td>Cotton</td>
<td>RCH-2</td>
<td>4500</td>
<td>5200</td>
<td>5150</td>
<td>12/21/2018</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Kheda</td>
<td>Balasinor</td>
<td>Cotton</td>
<td>RCH-2</td>
<td>4500</td>
<td>5200</td>
<td>5150</td>
<td>12/20/2018</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Kheda</td>
<td>Balasinor</td>
<td>Cotton</td>
<td>RCH-2</td>
<td>4700</td>
<td>5250</td>
<td>5150</td>
<td>12/19/2018</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Kheda</td>
<td>Balasinor</td>
<td>Cotton</td>
<td>RCH-2</td>
<td>4500</td>
<td>5250</td>
<td>5150</td>
<td>12/18/2018</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Kheda</td>
<td>Balasinor</td>
<td>Cotton</td>
<td>RCH-2</td>
<td>4500</td>
<td>5250</td>
<td>5150</td>
<td>12/17/2018</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Kheda</td>
<td>Balasinor</td>
<td>Cotton</td>
<td>RCH-2</td>
<td>4700</td>
<td>5250</td>
<td>5150</td>
<td>12/16/2018</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Kheda</td>
<td>Balasinor</td>
<td>Cotton</td>
<td>RCH-2</td>
<td>4800</td>
<td>5150</td>
<td>5150</td>
<td>12/15/2018</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.2. Pre-processing Dataset

Firstly, pre-processed data a bit so that it’s in a format and provide into a neural network. The process includes:

1. Remove irrelevant characters ("#$%&()*+,- /:;<=>?@[\]^\_`{|}~).
2. Convert all letters to lowercase (HeLlO -> hello).
3. As this is character level embedding, consider every character as a separate token.
4. Tokenize words
5. Standardize input length with padding

3.3. Develop A Machine Learning Model

The Auto Regressive Integrated Moving Average (ARIMA) is generally a model which leads to define the time series depends on the previous values. Specifically, the previous values are the lags and lagged prediction errors. Therefore, this ARIMA is used to predict the future values.

ARIMA models can be used to simulate any ‘non-seasonal’ time series that has patterns and is not random white noise. Figure 1 shows approach to implement the price prediction. Basic Idea of machine learning model to predict the Real-time Market Movement a trained model which keeps on improving the accuracy with time is implemented on backend. A Django based rest-API which takes care of data transmission between Backend and Frontend.

3.4 ARIMA and Seasonal ARIMA

The following is the general procedure for ARIMA models:

- Make the time series data stationary
- Construct the ARIMA Model or Seasonal ARIMA based on the data
- Use the model to make predictions

Auto Regressive Model describe as

\[ yt=c+ \varphi_1yt-1+ \varphi_2yt-2+\cdots+ \varphi pyt-p+ \epsilon \]  

Where \( p \) is the order of term and some other constant. Figure 2 shows that the time series data of Balasinor market. Crop price is very high on specific date and suddenly very low on other dates so which clearly indicate that data is not stationary.
To convert time series data into the stationary time series by simply differencing the series such step is called as Integration.

\[ y_t = y_t - y_{t-1} \]  \hspace{1cm} (2)

Above basic formula will help to transform a trendy to stationary as shown in figure 3.

The very first entity of architecture is the traders, who will interact with the proposed first component, a Hybrid Application which works on any kind of web browser and phone (Android or iOS). Farmers can also get connected with the GUI interface and receive updated crop price information and interact with the buyer. The second module of the architecture is the Machine Learning/Deep Learning model to generate the dynamic price prediction of crops by analyzing past data. Figure 4 shows that how information will flow across the different components and generate the analysis.

4. SIMULATION RESULTS AND DISCUSSION

Traders have to integrate bits of information regarding supply and demand and compare the guesses of fellow traders before he pursues his estimation by promising to buy or sell i.e., what are the current market trends? And what are the future predictions? Figure 5 shows the hybrid application page through which traders can interact and list of features available with them.

Once traders consent letter approved from the admin then only it’s allowed to view dashboard with market analysis for the cotton crops as shown in Figure 6. Before the implementing the model on the Balasinor market cotton RCH-2 variety model price shown in Figure 7. After the implementing model on same market dataset with model price (valid) compare with predicted price shown in Figure 8. So predicted price is more similar to valid price of cotton crops with minimum root mean square error was observed.
Table 2 shows the Balasinor (Mahisagar district of Gujarat, India) market prediction with compare to other nearest market price for cotton crops per quantal.

<table>
<thead>
<tr>
<th>Sl. No.</th>
<th>Market Name</th>
<th>Price (per Quintal)</th>
<th>Predicted Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Balasinor</td>
<td>4953</td>
<td>8/5/2021</td>
</tr>
<tr>
<td>2</td>
<td>Navsari</td>
<td>5426</td>
<td>8/5/2021</td>
</tr>
<tr>
<td>3</td>
<td>Amirgadh</td>
<td>5444</td>
<td>8/5/2021</td>
</tr>
<tr>
<td>4</td>
<td>Junagadh</td>
<td>5456</td>
<td>8/5/2021</td>
</tr>
<tr>
<td>5</td>
<td>Amreli</td>
<td>5469</td>
<td>8/5/2021</td>
</tr>
<tr>
<td>6</td>
<td>Rajula</td>
<td>5489</td>
<td>8/5/2021</td>
</tr>
<tr>
<td>7</td>
<td>Kadi</td>
<td>5491</td>
<td>8/5/2021</td>
</tr>
<tr>
<td>8</td>
<td>Mehsana</td>
<td>5499</td>
<td>8/5/2021</td>
</tr>
<tr>
<td>9</td>
<td>Anjar</td>
<td>5608</td>
<td>8/5/2021</td>
</tr>
<tr>
<td>10</td>
<td>Halvad</td>
<td>5662</td>
<td>8/5/2021</td>
</tr>
</tbody>
</table>

5. CONCLUSION
This solution is to build an Artificial Intelligence based Model which can predict cotton trends by analyzing the past year’s data (/trends). Approximate, Cotton trends can be predicted and thus it helps to grow the business as well as in bringing up India’s’ Economy.

REFERENCES


