© June 2021| IJIRT | Volume 8 Issue 1 | ISSN: 2349-6002

Krishi: A Price Prediction System

Chandni Agrawal[1], Piyush Kumar[2], Vijay Sehrawat[3], Naman Mishra[4], Tarachand Verma[5]
[1,2,3,4,5] *Raj Kumar Goel Institute of Technology/AKTU*

***Abstract* - In today's world where price prediction systems have become an essential need and are implemented in almost all sectors including estimation of transportation costs or real estate fares, for a country where a major area of land is being used for cultivation this paper tries to understand the importance of price prediction in this field and by overcoming the obstacles and hurdles, find a suitable solution by using the technologies and researches done till date to provide a farmer-friendly solution. This work was done keeping in mind the new Indian Agriculture Acts of 2020 as the middlemen or "Arhatis" (in Hindi) were not just a product link between farmer and customers but also the economic link and were often the ones that decided the prices of crops. With their absence an efficient system is required that can predict the prices for farmers and help them decide accordingly their future ventures.**

***Index Terms* - Arhati, Crop, Farmer, Machine Learning, MSP, SARIMA, Weather.**

I.INTRODUCTION

On market analysis, it was observed that today's market is not just dependent on the amount of crop cultivation but also on various other factors like weather and availability under a certain seasonal circumstance. Farmers lack ideas and information about various other alternatives. It was also observed that farmers instead of effectively utilizing their fields through methods like crop rotation, stuck on growing double seasonal crop throughout. The main reason for this was a fear of economic loss in experimenting with new crops. An alternative solution for this issue was found in the price prediction system. What if, the farmers were given a pre-determinant idea of the approximate price of a particular crop in the upcoming time, this would provide them a brief insight into the alternatives available in the market and help them plan their resources accordingly. This can prove to be invaluable agricultural information.

Agricultural information is essential for improving agricultural production. Specifically, agricultural productivity can arguably be improved by relevant, reliable, and useful information and knowledge. Agricultural information interacts with and influences agricultural productivity in a variety of ways. It can help inform decisions regarding land, labor, livestock, capital, and management. Agricultural productivity can arguably be improved by relevant, reliable, and useful information and knowledge. [1]

Another important step in providing an effective output to a problem is imparting the output information to the person the information is meant for and in an effective manner such that it can be presented and processed by the consumer. The problems arise when this customer base consists of consumers with approx. 19,500 dialects [2] it becomes difficult to design a common interface that can act similarly to provide and understand the input of all the customers and respond with equal efficiency.

II.LITERATURE REVIEW

Though there are several research institutions, universities, public offices, and libraries that have a number of studies, a huge amount of data, and chunks of information to impart agricultural knowledge, yet only an ounce of it is accessible by the farmers. This lack of information to the farmers of their products often leads them to suffer a loss. Thus, a farmer must know the actual price of his crops and should not be easily deceived by anyone.

On further analysis of recent researches of pertaining systems, we found that due to the diversity of the languages in Nigeria, it is presumed that for farmers to have access to agricultural information through the radio and television, the language of presentation has to be based on that of the listeners. [3] On implementing the same results in India, the feasibility of the product was questioned, as providing a language-based solution in such a diverse country was not possible. Therefore to overcome this barrier and to avoid any cof4nfusion and misunderstandings due to the varying languages in various parts of the country, we approached different researches that had encountered similar issues where we found an effective way to represent data that makes it more comprehensive. The use of images and presenting data in the pictorial form not only makes the data quite understandable but also user-friendly. [4]

The third broad aspect was our consideration of various price prediction models currently persisting in the market. Their evaluative study presented various deficits in this agricultural field regarding a variety of soils, and a further broader aspect of the analysis of crops shows that even different categories of the same crops may require varying seasonal requirements. For example - Price forecasting is more acute with vegetable crops particularly tomatoes due to its highly perishable nature and seasonality. The tomato is grown in practically every country of the world and is one of the most important agricultural products among fresh vegetables in most countries in the world. Turkey is among the countries producing various kinds of vegetables at a high production level due to suitable ecological conditions. [5]

III.ISSUES IN CURRENT APPLICATIONS

To better find conclusive solutions it is important to summarize the current problems at hand that persisted in the current systems

Models: The data which the present model needs to predict the prices call for input such as soil quality, weather, water content, etc. which requires extensive research on each land where the crop is grown. Such a task on large scale is not possible rendering such price prediction systems financially unattractive. Government (agmarknet.gov.in) provides only two types of data about any crop: price, and date of a crop from a particular mandi, thus an economical price prediction system needs to work on only these two input-dataset.

Language: Languages can often be a confusing mode of communication due to their skeptical nature, common reference of this disorientation in meanings can be observed where a common Indian vegetable "Baingan" (in Hindi) is translated as both Brinjal and Eggplant in English. Thus, the same meaning for different words can prove to be a major default on an input-based system.

Education: The coherent nature of a solution plays a major role in defining its success. A solution that cannot be understood by the consumer section is equivalent to null, thus in this regard low literacy rate of the farmers, the highly complex User Interface of the govt systems, websites, and other farmer portals have been facilitation-based system rather than application- based i.e., they have their main concern in providing the specified services rather than the simplicity or usability of the product.

IV.INTRODUCING IMAGE-BASED RECOGNITION SYSTEM

A good picture is capable of conveying a thousand words, we can show a picture of an object like a ball to two people of different language origin and both of them would be able to understand the message. A picture can in effect ground the meaning to an object or concept in the real world and act as a convenient bridge over language barriers. Thus, adopting a similar approach used in different research and implying it for our problem statement, we have a simple input idea structure, we will be using pictures as the user input method. The users should be able to input the vegetable, crop, or fruit they want the price of, and the interface after accepting the input and implementing image recognition post-cleaning would predict the price. [6]

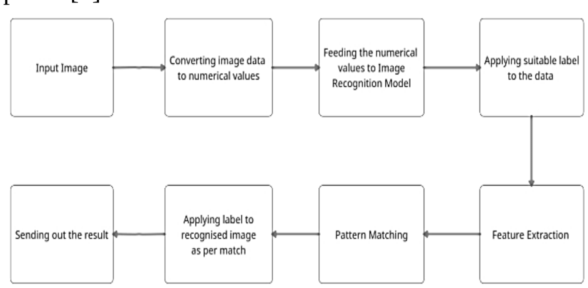

Fig 1: Data flow for image recognition

V.MACHINE LEARNING MODEL

Autoregressive Integrated Moving Average, or ARIMA, is one of the most widely used forecasting methods for univariate time series data forecasting. Although the method can handle data with a trend, it does not support time series with a seasonal component. As its name suggests, it supports both autoregressive and moving average elements. The integrated element refers to differencing allowing the method to support time-series data with a trend. ARIMA expects data that is either not seasonal or has the seasonal component removed, e.g. seasonally adjusted via methods such as seasonal differencing. [7]

A problem with ARIMA is that it does not support seasonal data. That is a time series with a repeating cycle. ARIMA expects data that is either not seasonal or has the seasonal component removed, e.g. seasonally adjusted via methods such as seasonal differencing.

An extension to ARIMA that supports the direct modeling of the seasonal component of the series is called SARIMA.

A seasonal ARIMA model is formed by including additional seasonal terms in the ARIMA. The seasonal part of the model consists of terms that are very similar to the non-seasonal components of the model, but they involve backshifts of the seasonal period.

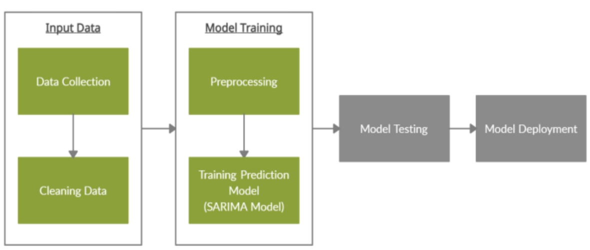

Fig 2: ML Model Transition

Seasonality usually causes the series to be non-stationary because the average values at some particular times within the seasonal span (months, for example) may be different than the average values at other times. Seasonal differencing is defined as a difference between a value and a value with lag that is a multiple of S. Seasonal differencing removes seasonal trend. [8]

SARIMA is an extension of ARIMA that explicitly supports univariate time series data with a seasonal component. It adds three new hyperparameters to specify the autoregression (AR), differencing (I), and moving average (MA) for the seasonal component of the series, as well as an additional parameter for the period of the seasonality.

Trend Elements:

Three trend elements require configuration.

p: Trend autoregression order.

d: Trend difference order.

q: Trend moving average order.

Seasonal Elements:

There are four seasonal elements.

P: Seasonal autoregressive order.

D: Seasonal difference order.

Q: Seasonal moving average order.

m: The number of time steps for a single seasonal period.

Together, the notation for a SARIMA model is specified as: SARIMA(p,d,q)(P,D,Q)m[]

VI.RESULT

Dataset Sourced from: agmarknet.gov.in

District	Udham Singh Nagar
Market Name	Kiccha
Commodity	Wheat
Date Range	06/04/2017 - 31/12/2019
Dataset Training-Test ratio	90: 10

Table 1: Input Dataset Source and training-test Distribution

Mean Squared Error of the forecasts	809.52
Mean Squared Percentage Error of forecasts	1.35%

Table 2: Machine Learning model Prediction's error percentage

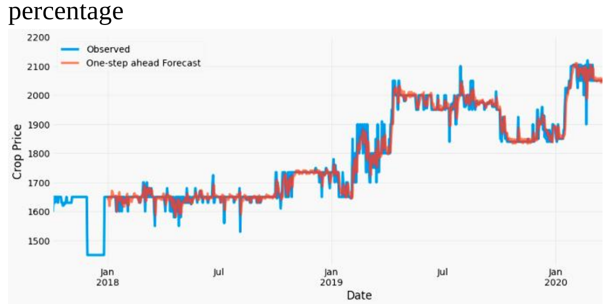

Fig 3: Observed & One Step Ahead Forecast Comparison of trained ML Model

VII.CALL BASED CHAT-BOT IMPLEMENTATION

Another additional feature that can be an addition to the solution in this approach can be the inclusion of the problem of remote connectivity of farmers. Many farmers situated in remote areas don't have access to internet facility for outsourcing, to solve this problem the solution includes the addition of a chatbot facility through Dialogflow by GCP using NLP and Google API, the chatbot will be configured to respond according to the simplistic requirement of the farmer reading the name of vegetable from the farmer from text/voice-based input medium. The voice-based medium would be configured via call; this will be using the default Dialogflow's call engine responding according to the variance in the input. Primarily the voice being configured will be set upon keeping in mind two languages i.e., Hindi and English. [10]

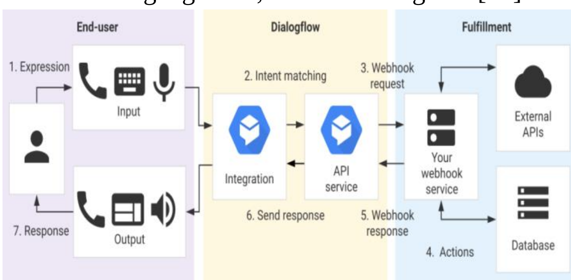

Fig 4: Dialog flow data transition

VIII.DRAWBACKS IN THE CURRENT APPROACH

Hardware dependent: The input medium that is image is dependent on the hardware of the device the image has been clicked from, thus an image from a lower configuration device may have varied or wrong result parameters for input.

Background Differentiation in Image: The image must be differentiated from the background for image recognition algorithms to work properly. If the algorithm fails to differentiate the image from the background, results are expected to be biased.

IX.CONCLUSION

The project work introduces an efficient crop price prediction system using the Seasonal Autoregressive Integrated Moving Average model. This model is capable of predicting the price even in absence of multiple parameters. The web-based approach would help in simplifying the interface as the farmer does not need to download additional applications. The system is scalable as it can be used to test different crops, fruits, and vegetables. Not only this, but the change in season is also taken into consideration. From the pictorial representation used in the project, we can eradicate the language barrier that hinders the communication between farmers and the technological interface. The images make the platform interactive and user-friendly. The solution also contains a chatbot that solves the connectivity issues and makes the platform remotely accessible.

REFERENCES

[1] Journal of Developing Country Studies, Factors That Influence Effective Communication of Agricultural Information among Farmers – The Case of Farmers in Southwest Kisumu Ward, Kisumu County, Miriam A. Awili, Prof. White, Dr. Stephen Kimotho, Pg-2
[2] More than 19,500 mother tongues spoken in India: Census, The Indian Express Article
[3] Journal of Developing Country Studies, Factors That Influence Effective Communication of Agricultural Information among Farmers – The Case of Farmers in Southwest Kisumu Ward, Kisumu County, Miriam A. Awili, Prof. White, Dr. Stephen Kimotho, Pg-3
[4] Article: Speaking Louder than Words with Pictures Across Languages Andrew Finch, Wei Song, Kumiko Tanaka-Ishii, Eiichiro Sumita Pg 34
[5] An analysis of tomato prices at wholesale level in Turkey: an application of SARIMA model. Adanacioglu, H; Yercanm M., ISSN 1808-2882 Pg 52
[6] Article: Speaking Louder than Words with Pictures Across Languages Andrew Finch, Wei Song, Kumiko Tanaka-Ishii, Eiichiro Sumita Pg 32
[7] Agricultural production output prediction using Supervised Machine Learning techniques, 2017 1st International Conference on Next Generation Computing Applications (NextComp), Md. Tahmid Shakoor; Karishma Rahman; Sumaiya Nasrin Rayta; Amitabha Chakrabarty
[8] An analysis of tomato prices at wholesale level in Turkey: an application of SARIMA model. Adanacioglu, H; Yercanm M.
[9] A Gentle Introduction to SARIMA for Time Series Forecasting in Python by Jason Browniee
[10] Dialogflow Fulfillment, Documentation, Google Cloud Platform