North South University

Crop Price Prediction using Machine Learning

Nafisa Ayman Prome 2012027042

Zia Uddin Chowdhury 2013754042

Juhaied Hossen 1921556642

Supervisor: Dr. Sifat Momen

A report submitted in partial fulfilment of the requirements
for the degree of Bachelors in Computer Science

in the

Department of Electrical and Computer Engineering

November 5, 2023

Declaration

All sentences or passages quoted in this document from other people's work have been specifically acknowledged by clear cross-referencing to author, work and page(s). Any illustrations that are not the work of the author of this report have been used with the explicit permission of the originator and are specifically acknowledged. We understand that failure to do this amounts to plagiarism and will be considered grounds for failure.

Name: Nafisa Ayman Prome

Signature:

Date:

Name: Zia Uddin Chowdhury

Signature:

Date:

Name: Juhaied Hossen

Signature:

Date:

Abstract

Our project emphasises the critical task of crop price prediction, which is required for informed agricultural practices. We have utilize three powerful algorithms individually and as an ensemble classifier: Decision Tree, Random Forest, and XGBoost. Our methodology entails training models on large datasets from the "Yearbook of Agricultural Statistics, Bangladesh"[1], spanning the years 2019, 2020, and 20211, with regularisation techniques used for fine-tuning. The experimental results show that the ensemble classifier outperforms individual models, with an impressive R2 score of 0.96. The novel aspect of the work is the strategic integration of various algorithms, which results in a comprehensive solution for accurate crop price prediction in a dynamic agricultural landscape offering essential assistance to farmers and other agricultural stakeholders.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Bangladesh is an agricultural country. Since 2000, Bangladesh's rural economy, particularly its agricultural sector has been a major force in the nation's attempts to reduce poverty. More than 70 percent of Bangladesh's population and 77 percent of its workforce live in rural areas. Nearly half of Bangladesh's workers and two-thirds in rural areas are directly employed by agriculture, and about 87 percent of rural households rely on agriculture for at least part of their income [2]. According to the provisional calculation of BBS, the contribution of agriculture to the GDP in FY 2021-22 is about 11.50 percent [3]. Despite farmers' significant contribution to the sectors, they do not receive a fair portion of the market prices for their goods because of weak market infrastructure. The report, "Dysfunctional Horticulture Value Chains and the Need for Modern Marketing Infrastructure: The Case of Bangladesh" also notes that producers get less than 40 percent of prices at the consumer level. Merchants get 43 percent while the remaining 17 percent is spent on transportation, preservation, sorting, packaging, and other expenses [4] We can see from the statistics that agriculture contributes significantly to Bangladesh's GDP, but the majority of farmers who contribute significantly to the economy continue to be exploited.

A crop price prediction system will empower farmers by providing accurate market pricing, reducing information asymmetry, and preventing exploitation by enabling informed decisions in selling their produce at fair rates. Farmers are being exploited on a daily basis due to a lack of knowledge and information. To address this issue we have come up with e Crop Price Prediction System using Machine learning techniques, a system that will help the country's agricultural development by giving farmers access to accurate crop prices. Our approach to addressing crop price prediction involves utilizing advanced machine learning algorithms incorporating key factors such as historical prices, pesticide usage, and climate data.

In the past, techniques like historical averages, supply and demand analysis, professional judgment, and on-the-ground market observations were used to predict crop prices. These methods were helpful, but they lacked the accuracy and flexibility of more recent methods that use advanced algorithms for more accurate results. Our approach, in contrast to previous approaches, makes use of smart programs that quickly learn and adapt, taking into account the latest information such as market trends and weather conditions. This improves prediction precision and benefits farmers. Furthermore, farmers don't need to be tech experts to use it because it's designed to be simple to use.

Our goal is to create a system that has the potential to accurately predict upcoming crop prices, offering essential assistance to farmers and other agricultural stakeholders. The paper's novelty is its high-accuracy machine-learning approach to crop price prediction, which will benefit thousands of farmers.

This report is further organized into right chapters including Chapter 3: Literature Review, Chapter 4: Methodology, Chapter 5: Approach, Chapter 6: Results and Analysis, Chapter 7: Discussion, Chapter 8: Limitations, Chapter 9: Ethical Issues and Chapter 10: Conclusion

Chapter 3 (Literature Review): In this chapter, several approaches, models, and conclusions pertaining to crop price prediction are highlighted as part of an extensive analysis of previous research.

Chapter 4 (Methodology): The research technique is covered in detail in this section. It describes the data sources, methods of collecting, and stages involved in data preparation. This chapter provides a thorough explanation of the machine learning model's training, testing and deployment.

Chapter 5 (Approach): Building upon the methodology, this chapter elaborates on the approach taken in creating the predictive model. It explains in details about Decision Tree, Random Forest, XGBoost along with the formation of an Ensemble Regressor by combining them. Furthermore, it also explores Elastic Net Regression.

Chapter 6 (Results and Analysis): This section provides a thorough presentation of the study's results. It offers a detailed analysis of the model's performance, including metrics such as MSE, RMSE, and R-squared. Additionally, it discusses any observed patterns or insights gleaned from the predictions.

Chapter 7 (Discussion): The discussion highlights the strengths of the project, emphasizing its ensemble modeling approach, unique dataset source, and user-friendly interface, distinguishing it from previous work and offering accurate crop price predictions for agricultural stakeholders.

Chapter 8 (Limitations): This chapter outlines potential limitations of the project, discussing factors which may impact prediction accuracy.

Chapter 9 (Ethical Issues): This chapter addresses ethical considerations related to data consent, ownership, privacy, fairness, environmental impact, and software transparency in the project, highlighting the steps taken to address and mitigate these concerns.

Chapter 10 (Conclusion): This last chapter provides a succinct overview of the main conclusions and their implications from the study. Along with outlining helpful suggestions, it also points out directions for future research and emphasizes the importance of the study in relation to crop price forecast and the agriculture industry.

Chapter 2

Literature Review

2.1 Forecasting Crop Price Using Various Machine Learning Approaches

Chaitra.B et al. proposed Forecasting Crop Price Using Various Machine Learning Approaches [5] where they used Time Series Algorithms such as ARIMA, SARIMA, and NARX, as well as Decision Trees combining rainfall and temperature parameters, for precise price predictions. Furthermore, Neural Network algorithms, such as CNN and LSTM, contribute to precise forecasting by analyzing multiple crops on different time scales, with performance measured using metrics like MAPE and RMSE. In beetroot price prediction using SARIMA with previous year data, a high accuracy of 93% was achieved. For green chilly, employing the LM algorithm with FFNN and the previous year's price data yielded an impressive accuracy of 99.95%, showcasing the model's effectiveness.

2.2 A Methodology for Crop Price Prediction Using Machine Learning

G. Thapaswini et al. proposed A Methodology for Crop Price Prediction Using Machine Learning [6], which offers a comprehensive analysis of machine learning algorithms in agriculture, employing decision tree and neuro-evolutionary algorithms. Neuro-evolutionary algorithms optimize neural network structures and parameters through evolutionary principles for solving complex problems. The proposed framework is evaluated with a sample dataset from Kaggle, demonstrating enhanced crop productivity and pricing estimation, achieving a precision rate of 86%, a recall rate of 87%, and an F1 score of 86.5% in crop yield predictions.

2.3 Enhanced Crop Price Prediction and Forecasting System

S. Sajithabanu1 et al. proposed a model to show an Enhanced Crop Price Prediction Forecasting System [7]. This research explores agricultural pricing prediction, focusing on rice, wheat, milk, and fruit categories like mangoes. Utilizing a dataset from NAFHA, spanning 5 years of past prices, the study employs Naive Bayes for clustering algorithms, achieving an 85% accuracy in predictions. Additionally, the deployment of the Random

Forest Algorithm, considering dynamic environmental and climatic data, demonstrates its suitability for handling dynamic raw data, ensuring a robust and accurate prediction model. The implementation process involves using Decision Trees and Support Vector Machines to process dynamic raw data, contributing to the feasibility and accuracy of the prediction system.

2.4 FarmEasy: An Intelligent Platform to Empower Crops Prediction and Crops Marketing

Md Eshak et al. proposed about - FarmEasy: An Intelligent Platform to Empower Crops Prediction and Crops Marketing [8]. The authors suggest developing an intelligent system that can anticipate the best crops to plant depending on a farmer's location, provide in-depth advice on everything from soil preparation to crop production, and organize the entire crop marketing chain from farmers to consumers. The paper integrates real-time climate, meteorological, and soil data and uses Voting Regressor which offers a novel approach to crop yield prediction by averaging the predictions from Random Forest Regression (RFR) and Support Vector Regression (SVR) models. The authors retrieved information from the "Year Book of Agricultural Statistics of Bangladesh" and have compiled the data. For the period of 2013–2019, data on the soil, weather, and agricultural yields in 64 districts of Bangladesh were collected using the yearbooks of 2015, 2016, 2017, 2018, and 2019. But the paper also has its limitations as the authors have only focused on 6 types of crops due to the shortage of data. Additionally, if a field survey were conducted rather than using a pre-made dataset, the credibility and contribution would have been higher.

2.5 Crop Yield Prediction using Machine Learning Techniques

Ramesh Medar et al. have proposed Crop Yield Prediction using Machine Learning Techniques [9]. Agriculture's vital role in India's economy hinges on effective crop selection. Machine learning methods, including traditional statistics, optimize crop production by considering factors like weather, soil type, and geography. The Naive Bayes method and K-Nearest Neighbour method were employed with a Java application to predict crop yield rates. The accuracy of both methods was assessed for crop selection, aiding farmers in optimizing their crop choices. Even though a dataset is used but the authors didn't provide any source of their data. The authors have come to the conclusion that their model can be used to facilitate crop selection and yield prediction, ultimately contributing to improved agricultural performance and economic growth.

2.6 An intelligent Crop Price Prediction using suitable Machine Learning Algorithm

Ishita Ghutake et al. have proposed An intelligent Crop Price Prediction using suitable Machine Learning Algorithm [10].It has been a challenging challenge for farmers to plan their crops for the upcoming season because it is difficult to forecast the metrics of pricing that their produce would fetch in a given season, which will typically be depending on dynamic weather conditions.They developed the Crop Price Prediction Website for crop forecasting where they took data from the government of more than 20 crops and represented the data in a structured manner representing the increase and decrease in prices of crops per month and further showing the crops details like its type, location, and export factors for the ease of the farmers to plan and manage their finances and sown/harvested in a timely manner.To establish a refined platform for interaction that farmers could use and use to make predictions for the upcoming 12 months, they conducted an in-depth statistical examination of historical data. Features of the dataset was month ,year and rainfall. They trained their model using decision tree ,random forest and linear regression. Their model has an accuracy of 92%, with monthly variations. They didn't use the temperature,humidity attribute which have a significant impact on crop price prediction so this is one of the limitations of this model.

2.7 Meta-Learning Based Adaptive Crop Price Prediction for Agriculture Application

Dhanasekaran K et al. have proposed a Meta-Learning Based Adaptive Crop Price Prediction for Agriculture Application [11]. For time-series prediction, a novel method is put forth that makes use of the ability to anticipate crop yields and prices for a small number of crops in order to pinpoint the pertinent data concerning market prices and crop yields. Several risky circumstances, including changes in the climate, variations in the sea. This work has used climate data collected from the National Climate Data Center (NCDC) for calibration of the system.They used meta learning technique.Meta-learning based neural network uses 20 hidden layers and 50 neurons in each layer. .This method establishes the linkages between price and agricultural yield component relationships and eliminates less important characteristics to improve performance.The suggested approach outperforms other models utilized for comparison, including SOM and LSTM, in terms of performance.

2.8 Vegetable Price Prediction Based on PSO-BP Neural Network

Y. E. Lu and colleagues introduced an innovative approach in their paper titled "Vegetable Price Prediction Based on PSO-BP Neural Network" [12]. Their study aimed to provide precise forecasts of vegetable prices by harnessing the power of data and cutting-edge technology.The researchers used a global stochastic optimization method called Particle Swarm Optimization (PSO) to adjust the parameters of a Backpropagation (BP) neural network by taking advantage of the dynamic nature of price changes.This model is unique in that it can surmount two significant obstacles that are frequently encountered in predictive modeling: local minima and over-fitting. The model produced amazing results by applying the PSO method to optimize the BP neural network's initial weights and thresholds. It outperformed conventional BP approaches in terms of performance and decreased training errors significantly.In essence, the PSO-BP model presented by Lu and his team is not just an enhancement; it's a game-changer in the field of vegetable price prediction, offering a more accurate and reliable tool for understanding and anticipating price dynamics in the market.

2.9 Data Analytics in Farming: Rice Price Prediction in Andhra Pradesh

Pundru Chandra Shaker Reddy et al. have proposed DATA ANALYTICS IN FARMING: RICE PRICE PREDICTION IN ANDHRA PRADESH using suitable Machine Learning Algorithm [13] .Farmers, particularly those who reside in rural areas, greatly benefit from accurate price prediction of crop production. Although a lot of study is being done in this field, accuracy still has to be improved.The climatically method of Holt-Winter and ARIMA is a commonly used numerical technique for predicting information with seasonality and flair. The dataset was obtained from the Agriculture division of Andhra Pradesh.It included the price of rice for each month from 2002 to 2021 for 13 stations in Andhra Pradesh. The training set consisted of perceptions from 2001 to 2015, whereas the attributes from 2016 to 2020 were used for testing.They use LSTM as a model because it can detect the information's non-linear dependence, making it ideal for time-series prediction. Their model performed really well in LSTM.

2.10 An Enhanced Approach for Crop Yield Prediction System Using Linear Support Vector Machine Model

Farmers receive assistance from K. Priyadharshini[14] in determining if a particular crop is appropriate for a given season and crop price. The author employed a variety of prediction algorithms, including decision trees in machine learning, SVM, KNN method, and linear regression. This work offers a novel approach to supply sufficient support vectors for a Support Vector Machine (SVM) classification based on auxiliary data. Precision categorization is used in a real-time agriculture application scenario using this refined technique is applied, improving production management. The suggested SVM approach provides better accuracy than the current framework. The author comes to the conclusion that this strategy can be adopted in a number of government sectors, including APMC and Kissan call center, etc., via which farmers and the government can obtain data regarding the market price and the anticipated crop yield in the future.

2.11 Crop Yield Prediction based on Feature Selection and Machine Learners: A Review

The study conducted by Girish K.[15] utilized monthly wholesale price series of soybean and rapeseed-mustard to demonstrate the superiority of ANN over linear model approaches. Empirical evidence demonstrates that ANN models are able to capture a significant number of directions of monthly price change when compared to linear models. Additionally, it has been found that when data exhibit nonlinear patterns, combining linear and nonlinear models produces forecasts that are more accurate than the models' individual performances. With the integration of linear and nonlinear forecasting techniques, the current study aims to design an ANN-based decision support system that is easy to use.

2.12 Notable Work Comparisons

Table 2.1: *Comparison of Notable Works in Crop Price Prediction*

Paper	Algorithms	Limitations	Key Findings	Our improvement
[5]	Time Series Algorithms such as ARIMA, SARIMA, and NARX	Lengthy literature reviews but reduced contribution of the authors	Comparative Analysis of different algorithms and their price prediction capability	We provided in-depth methodology explaining each step of our Crop price prediction Model
[6]	Naive Bayes, KNN and ANN	No credible source for the dataset	Improved crop yield predictions	We have used data from Yearbook of Agricultural Statistics, Bangladesh
[7]	Naive Bayes, Random Forest	Lower accuracy of 85% and no steps show to regularize or improve it	Accurate price predictions for multiple crops	Carried out Elastic Net Regularization
[8]	Voting Regressor - Random Forest and Support Vector	Focused only on 6 types of crops	Enhanced Crop Yield Predictions	Our model predicts the price of around 250 different crops
[9]	Naive Bayes, KNN, Decision Tree and SVM	The source of the dataset is not specified	Model can be used to facilitate crop selection and yield prediction	The source of the dataset is explicitly mentioned
[10]	Decision Tree, Random Forest and Linear Regression	Temperature and humidity not taken into account	92% accuracy in crop price prediction	We have taken temperature, humidity, and pesticide usage into account

All of the restrictions listed in the aforementioned table have been addressed, and we have also supplied a detailed methodology outlining each stage of our crop price prediction model. Our reliable source of data is the "Yearbook of Agricultural Statistics, Bangladesh", which is a credible source. Our algorithm can forecast the price of about 250 different crops, and we have included factors like temperature, humidity, and pesticide use because these are significant factors in crop price prediction. Additionally, we used Elastic Net Regularization.

Chapter 3

Methodology

The methodology involves data collection from the "Yearbook of Agricultural Statistics, Bangladesh" [1] between 2019 and 2021. Preprocessing and feature selection are executed, with emphasis on attributes like crop variation, weather parameters, and temporal factors. A robust Voting Regressor model, combining Random Forest, Decision Tree, and XGBoost, delivers high predictive accuracy with an R-squared score of 0.96. Model deployment through a Streamlit web application enables user-friendly access to real-time crop price predictions, catering to informed agricultural decision-making.

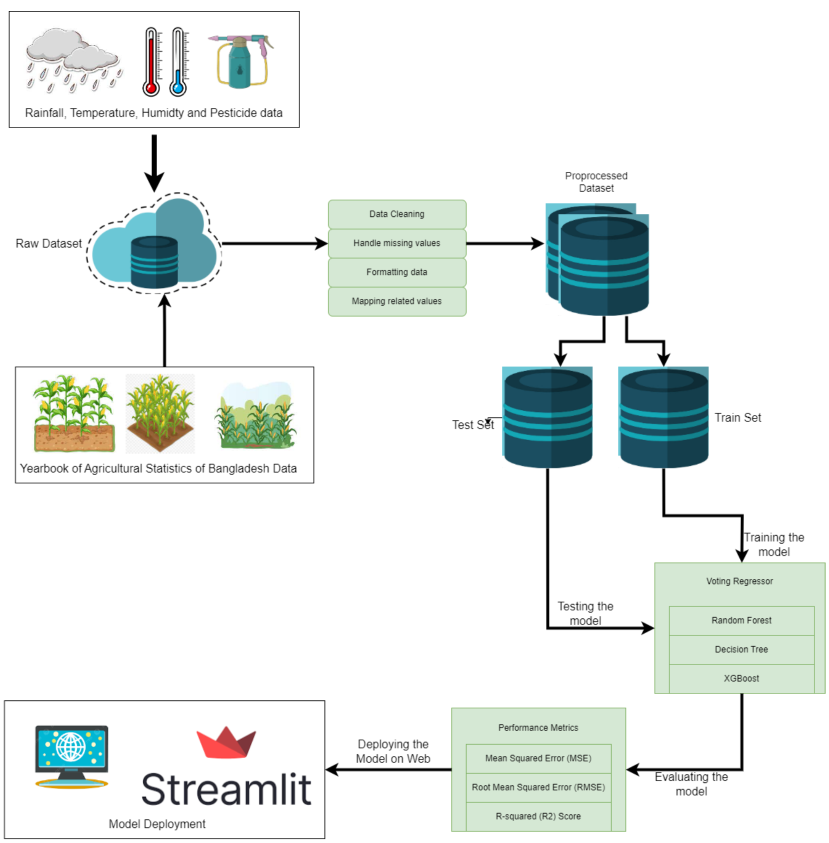

Figure 3.1: *Methodology*

3.1 Problem Formulation

The problem at hand revolves around forecasting crop prices, a critical concern for both farmers and consumers. Our goal is to create a reliable predictive model that uses environmental and historical data to make timely and accurate predictions of future crop prices. This will enable the agriculture sector to make well-informed decisions.

3.2 Data Collection

The dataset utilized for crop price prediction was sourced from the "Yearbook of Agricultural Statistics, Bangladesh", spanning the years 2019, 2020, and 2021. Extensive data were methodically collected from diverse districts across Bangladesh, providing a comprehensive picture of agricultural dynamics. Crop pricing, meteorological parameters (rainfall, temperature, humidity), and pesticide use were all taken into account. Monthly observations from January to December allowed for an in-depth examination of temporal patterns, while the addition of randomized numerical values increased the dataset's realism and variability. These data were all available in different tables, which were then manually integrated to create the dataset.

: Department of Agriculture Marketing

10.1 Monthly Average Wholesale Price-2021

(Per Quintal/Taka)

Sl no	Name of Comodity		Year-2020												
			January	February	March	April	May	June	July	August	September	October	November	December	Average
01	Paddy Aman	Fine	3006	3204	3307	3265	3238	3198	3268	3177	3203	2812	2988	3004	3139
02	,,	Medium	2780	2990	3064	2984	2962	2980	2958	2925	2780	2612	2574	2595	2850
03	,,	Coarse	2543	2732	2844	2762	2730	2691	2656	2620	2455	2260	2360	2398	2588
04	,,	Pajam	3450	-	-	-	2975	2975	-	-	-	-	-	3053	3113
05	"	shugandhi	4494	4340	4338	4267	4200	-	-	-	-	-	-	4654	4382
06	Paddy Aus	medium	2625	2850	2850	2850	2850	2850	2850	2479	2444	2485	2500	2487	2677
07	,,	coarse	2575	2625	2625	2625	2588	2625	2338	2360	2181	2225	2256	2298	2443
08	Paddy Boro	fine	3199	3327	3449	2783	2467	2668	2893	2961	2948	3005	2984	3084	2981
09	,,	medium	3005	3118	3223	2524	2355	2533	2663	2714	2727	2722	2763	2838	2765
10	"	coarse	2716	2823	2884	2187	2106	2320	2425	2458	2450	2415	2418	2498	2475
11	,,	pajam	-	-	-	2450	1975	2387	-	-	-	-	-	-	2271
12	Rice Aman	fine	5619	5756	5900	5908	5851	5781	5852	5843	5836	5774	5748	5817	5807
13	,,	medium	4682	4800	4910	4867	4783	4738	4750	4787	4820	4673	4614	4629	4755
14	,,	coarse	4056	4173	4256	4244	4175	4140	4156	4167	4151	4066	3947	3984	4126
15	,,	pajam	5480	5656	5648	5778	5733	5605	5646	5679	5521	5475	5630	5690	5628
16	Rice boro	fine	5608	5774	5880	5850	5608	5540	5650	5690	5618	5582	5601	5654	5671
17	,,	medium	4889	5014	5102	5022	4717	4675	4767	4811	4767	4711	4706	4745	4827
18	,,	coarse	4096	4183	4295	4268	4055	4049	4115	4151	4087	4024	3972	4012	4109
19	,,	pajam	4475	4267	4292	4450	4225	4781	4800	5367	4633	4800	5150	5170	4701
20	Rice OMS	course	4083	4131	4385	4339	4330	4300	4306	4375	4370	4375	3867	3840	4225
21	Aromatic rice	Chinigura	8663	8661	8651	8659	8575	8488	8519	8516	8527	8519	8532	8600	8576
22		Kalojira	8084	8065	8035	8067	8000	7973	7948	7911	7954	7979	7988	7979	7999
23		kataribhog	6539	6525	6631	6591	6645	6632	6629	6631	6730	6701	6688	6770	6643
24	Wheat local	red	2572	2613	2647	2617	2646	2630	2616	2619	2664	2828	2844	2873	2681
25	,,	white	2611	2670	2644	2601	2612	2620	2625	2606	2664	2766	2857	2937	2684
26	Imported	red	-	2650	2750	-	-	2750	2750	2750	2750	2750	-	-	2736
27	,,	white	2636	2690	2762	2680	2738	2721	2711	2727	2714	2840	2920	2966	2759
28	Ata	loose	2651	2712	2719	2694	2657	2644	2594	2633	2768	2908	2987	3055	2752
29	,,	Packet	3020	3077	3108	3107	3100	3096	3085	3108	3231	3318	3410	3525	3182
30	flour	loose	3193	3231	3232	3307	3234	3213	3210	3253	3408	3599	3734	3842	3371
31	,,	packet	4020	4045	4058	4102	4078	4063	4085	4107	4174	4266	4338	4485	4152
32	Maize	-	2022	2064	1986	1964	1967	2065	2212	2254	2413	2480	2545	2555	2211
33	Kalai	mash	8609	8625	8619	8584	8586	8578	8569	8424	9816	8291	8276	9438	8701
34	,,	mashur	6131	6144	6269	6025	6025	6025	6025	6000	6331	5972	6029	9717	6391
35	,,	mug	9186	9172	9180	9171	9156	9144	9156	8618	8553	8656	8666	8734	8949
36	,,	Chola	6837	6828	6800	6850	6835	6824	6832	6669	7257	6498	6746	6519	6791
37	,,	Motor local	5377	5182	5181	5116	5116	4953	5116	5063	5013	5132	5144	5225	5135
38	,,	Motor foreign	3550	3529	3559	3498	3515	3532	3522	3534	4397	3595	3595	3594	3618
39	,,	kheshari	5227	5254	5258	5195	5195	5193	5195	4709	5005	4661	4680	4698	5023
40	Dal	Mash gota	10668	10645	10754	11189	11110	11111	11103	11105	11050	11068	11057	11054	10993
41	,,	Broken	11326	11316	11293	11070	11017	11005	10964	11200	10949	10677	10786	10751	11030

Figure 3.2: *Year Book of Agricultural Statistics of Bangladesh*

	A	B	C	D	E	F	G	H	I	J	K	L
1	Name of Commodity	Variation	Year	Month	Rainfall	Temperature	Humidity	Pesticide	Price			
2	Paddy Aman	Fine	2020	January	27	27	83	4333	1994			
3	Paddy Aman	Medium	2020	January	45	25	76	4333	1753			
4	Paddy Aman	Coarse	2020	January	82	24	75	4333	1588			
5	Paddy Aman	Pajam	2020	January	56	24	86	4333	2236			
6	Paddy Aman	shugandhi	2020	January	13	27	87	4333	3988			
7	Paddy Aus	medium	2020	January	45	26	83	4333	2616			
8	Paddy Aus	coarse	2020	January	36	27	78	4333	2343			
9	Paddy Aus	pajam	2020	January	14	24	84	4333	1975			
10	Paddy Boro	fine	2020	January	21	23	85	4333	2342			
11	Paddy Boro	medium	2020	January	70	25	75	4333	1940			
12	Paddy Boro	coarse	2020	January	10	23	76	4333	1648			
13	Paddy Boro	pajam	2020	January	19	27	79	4333	2375			
14	Rice Aman	fine	2020	January	15	24	84	4333	4792			
15	Rice Aman	medium	2020	January	46	25	77	4333	3298			
16	Rice Aman	coarse	2020	January	13	27	70	4333	2697			
17	Rice Aman	pajam	2020	January	47	24	71	4333	4551			
18	Rice Aus	Medium	2020	January	64	24	86	4333	0			
19	Rice Aus	coarse	2020	January	40	27	83	4333	2661			
20	Rice Aus	Pajam	2020	January	14	26	70	4333	0			
21	Rice boro	fine	2020	January	60	25	78	4333	4557			
22	Rice boro	medium	2020	January	40	24	86	4333	3474			

Figure 3.3: *Raw Dataset*

3.3 Data Preprocessing

Since the dataset consisted of categorical data such as name of the month, name and variation of the crop, we had to convert these to numeric data by creating a map to convert the categorical values into numeric representations. We used an ordered dictionary to maintain the original order of values, ensuring a unique numerical mapping for each category of data. So the model would only be trained on numerical values.

1	Name of Commodity	Variation	Year	Month	Rainfall	Temperature	Humidity	Pesticide	Price
2	0	0	0	0	27	27	83	4333	1994
3	0	1	0	0	45	25	76	4333	1753
4	0	2	0	0	82	24	75	4333	1588
5	0	3	0	0	56	24	86	4333	2236
6	0	4	0	0	13	27	87	4333	3988
7	1	5	0	0	45	26	83	4333	2616
8	1	6	0	0	36	27	78	4333	2343
9	1	7	0	0	14	24	84	4333	1975
10	2	8	0	0	21	23	85	4333	2342
11	2	5	0	0	70	25	75	4333	1940
12	2	6	0	0	10	23	76	4333	1648
13	2	7	0	0	19	27	79	4333	2375
14	3	8	0	0	15	24	84	4333	4792
15	3	5	0	0	46	25	77	4333	3298
16	3	6	0	0	13	27	70	4333	2697
17	3	7	0	0	47	24	71	4333	4551
18	4	1	0	0	64	24	86	4333	0
19	4	6	0	0	40	27	83	4333	2661
20	4	3	0	0	14	26	70	4333	0
21	5	8	0	0	60	25	78	4333	4557
22	5	5	0	0	40	24	86	4333	3474

Figure 3.4: *Preprocessed Dataset*

3.4 Feature Selection

The names of the commodities, variations, year, month, rainfall, temperature, humidity, and pesticide are the features of our datasets. Our model will forecast the price using this. by learning about the agriculture economics and the elements that historically affect crop pricing. We realized that each feature in our model is really important. However, temperature, humidity, rainfall, and month are marginally more significant than the other factors.

Name of the Commodity: The kind or name of the crop can be an important characteristic because various crops have distinct market prices, demand, and susceptibilities to environmental factors. It helps predict price fluctuations specific to each crop.

Variation: The ability to identify patterns, seasonal fluctuations, or market variables influencing crop prices makes crop price variation a valuable attribute.

Year and Month: Knowing the month and year is essential to seeing yearly and seasonal patterns in crop prices. They can help develop price change models based on monthly market trends, harvest times, and general economic conditions, among other variables.

Rainfall, temperature, and humidity are three aspects of the weather that have a direct bearing on crop quality and growth. They have an impact on agricultural prices in the end as well as supply and demand. A drought, for instance, could result in lower supplies and increased costs.

3.5 Splitting of Dataset

The dataset was partitioned into two distinct subsets: a training set and a testing set. This partitioning was performed using a common practice, adhering to a 70:30 ratio. In this configuration, 70% of the dataset was designated for training the machine learning model, while the remaining 30% was exclusively reserved for evaluating the model's performance during testing. This approach ensures a robust and unbiased assessment of the model's predictive capabilities, providing valuable insights into its effectiveness.

3.6 Model Training and Testing

In the pursuit of crafting a robust and highly accurate model for crop price prediction, an intricate and meticulous approach was employed. The centerpiece of this endeavor was the implementation of a Voting Regressor model, artfully amalgamating the formidable predictive capabilities of not one, but three distinguished algorithms: Random Forest, Decision Tree, and XGBoost. The result of extensive and iterative testing, this complex orchestration of numerous algorithms was not a coincidence. The strengths and weaknesses of each algorithm were determined by carefully examining, fine-tuning, and rigorous evaluation. After a number of painstaking iterations, the best combination was found, and this led to the creation of a model that promises to be consistently accurate and stable when predicting crop prices. The extraordinary effectiveness of this model is derived from its capacity to combine the knowledge of three different algorithms, balancing their ability to forecast outcomes and yielding results that are unmatched in quality. A new benchmark of excellence has been set by this clever combination of approaches, guaranteeing that this model is ready to completely transform the crop price forecast industry.

```
# Create individual regression models
rf_model = RandomForestRegressor(n_estimators=100, random_state=42)
dt_model = DecisionTreeRegressor(random_state=42)
en_model = ElasticNet(alpha=best_alpha, random_state=42)

# create xgboost models
num_xgb_models = 50
xgb_models = [xgb.XGBRegressor(objective="reg:squarederror", random_state=42) for _ in range(num_xgb_models)]

# Create a Voting Regressor with the individual models
ensemble_model = VotingRegressor(
    estimators=[
        ("rf", rf_model),
        ("dt", dt_model),
        ("en", en_model),
    ] + [('xgb' + str(i), xgb_model) for i, xgb_model in enumerate(xgb_models)]
)

# Fit the ensemble model on the training data
ensemble_model.fit(X_train, y_train)

# Make predictions on the test data
y_pred = ensemble_model.predict(X_test)
```

Figure 3.5: *The Voting Regressor Model*

3.7 Model Evaluation

It is crucial to conduct a thorough performance evaluation of the model in order to create an excellent crop price prediction tool. Since we are dealing with a regression problem, the evaluation focuses on important metrics that show how well the model predicts the future. Mean Squared Error (MSE): The average squared difference (MSE) between the actual and forecast crop prices is a crucial measure. The model produces an MSE value of 15,115,290.16 in this situation. A smaller mean square error (MSE) indicates more predictive precision. This figure represents the magnitude of prediction mistakes.

Root Mean Squared Error (RMSE): RMSE is the square root of the MSE, and it serves as an intuitive measure of the model's prediction accuracy. The RMSE for this model is recorded at 3,887.84. A lower RMSE value implies that the model's predictions are closer to the true crop prices, underlining its commendable precision.

R-squared Score (R2): The statistical measure known as R-squared, or coefficient of determination, shows how much of the variance in crop prices can be explained by the model. Here, the R2 value rises to a very respectable 0.96. A model is considered to have excellent explanatory power if it can explain 96% of the variance in crop prices, as indicated by an R2 score of 0.96, which is equivalent to a perfect model.

When taken as a whole, these evaluation indicators offer strong proof of the Voting Regressor model's ability to predict crop prices. Its ability to create accurate forecasts is supported by the low MSE and RMSE values, and its ability to explain the underlying patterns and correlations in the data is demonstrated by the strong R2 score. Because of this, our model has the potential to have a big impact on crop price prediction by offering insightful information and assisting decision-makers.

3.8 Model Deployment

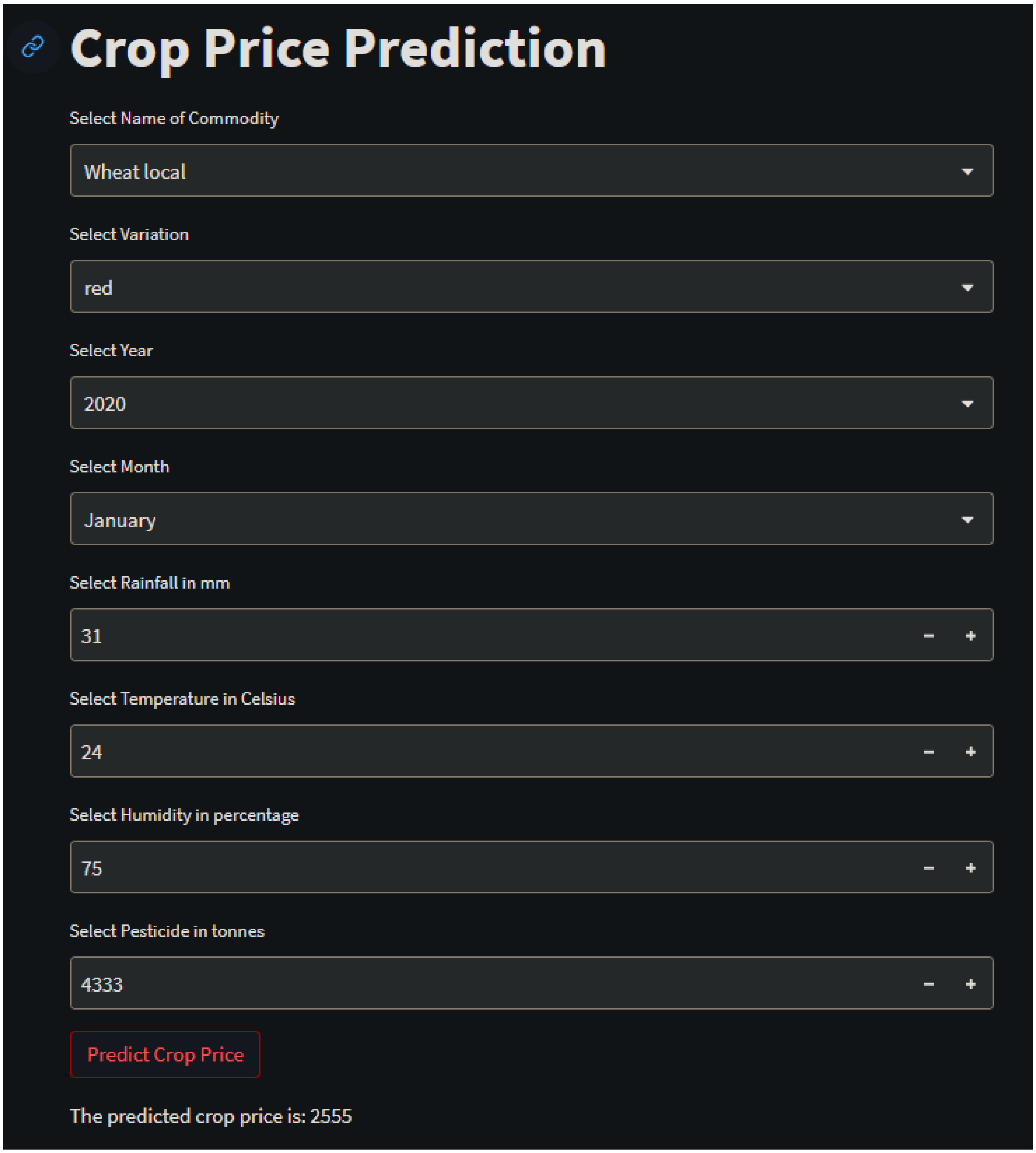

Figure 3.6: *Model Deployment*

The project entails the implementation of a machine learning model for crop price prediction, which is a multi-step procedure. First, the ‘joblib‘ library is used to export and preserve the machine learning model that has already been trained. Once saved as ”model.pkl,” this model is loaded with ease into the Streamlit web application. This crucial step ensures that the model’s predictive skills can be used quickly and effectively, removing the need for repeated model training.

The crop price prediction interface is then made user-friendly by integrating the model into a Streamlit web application. Users have the option to enter values for particular attributes, such as the commodity name, variation, year, and month, in addition to numerical parameters like humidity, temperature, rainfall, and pesticide levels. By selecting the ”Predict Crop Price” button, customers can initiate the prediction process after entering their desired numbers. Real-time predictions are made possible by the smooth communication between the loaded model and the user interface, giving users quick access to projected crop prices. It’s a crucial component of using web applications to implement machine learning models, which simplify user interfaces and provide accessibility to prediction tools.

Chapter 4

Approach

In this chapter, we will go over the architecture of our crop price prediction model in depth. We will discuss about the Exploratory Data Analysis and We are going to present outcomes of our various experiments.

4.1 Exploratory Data Analysis (EDA)

We examined crucial features such as Rainfall, Temperature, Humidity, and Pesticide in our Exploratory Data Analysis (EDA) covering 2019 to 2021. Scatter plots revealed correlations between these variables, leading our further feature selection and ensemble model decisions for accurate crop price predictions.

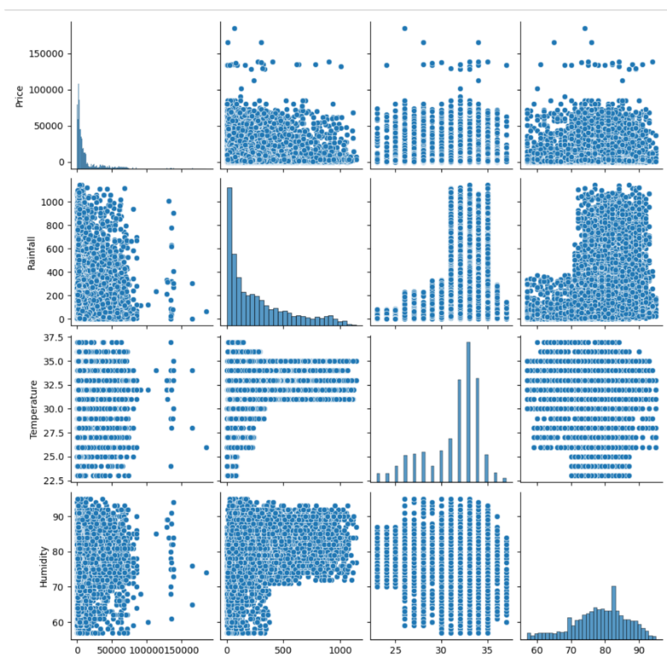

Figure 4.1: *Exploratory Data Analysis (EDA) Scatter Plot: Unveiling Relationships Among Key Agricultural Variables (2019-2021*

Here is the correlation Matrix of the dataset.

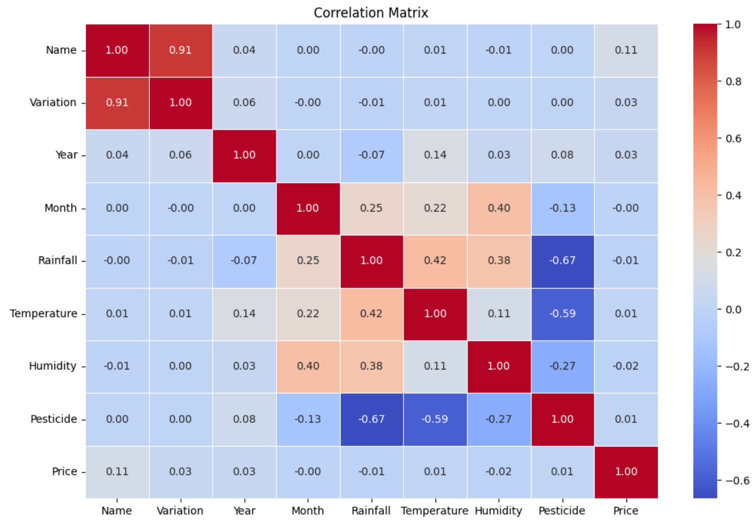

Figure 4.2: *Correlation Metrix*

4.2 Algorithms

We have used various machine learning algorithms in our crop price prediction model. We have used single Decision Tree, Random Forest also We have applied an ensemble classifier which incorporated Decision Tree, Random Forest, and XGBoost algorithms.

4.2.1 Decision Tree

A decision tree is a type of predictive modelling algorithm that creates a tree-like structure by recursively splitting the data according to feature requirements. This allows features to be mapped to outcomes. Every internal node represents a choice made in response to a feature, whereas every leaf node indicates a predicted result.

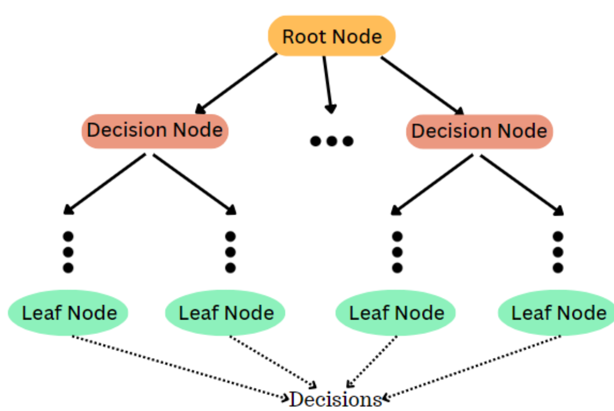

Figure 4.3: *A Decision Tree Algorithm.*

To begin our crop price prediction project, we first used a single Decision Tree Regressor on a carefully preprocessed dataset that included randomised temporal features for the years 2020 and 2021, such as temperature, humidity, and rainfall. After analysing the data we split it into training and testing sets, we used quantitative metrics including mean squared error (MSE), root mean square error (RMSE), and R-squared (R2) to assess the model's efficancy and it performed really well.

4.2.2 Random Forest

Random Forest is an ensemble learning method that builds many decision trees based on a random selection of features during training. The regression model is robust and accurate since it aggregates the predictions of individual trees. It is an adaptable option for machine learning algorithms because to its capacity to manage high-dimensional datasets, reduce overfitting, and capture intricate correlations.

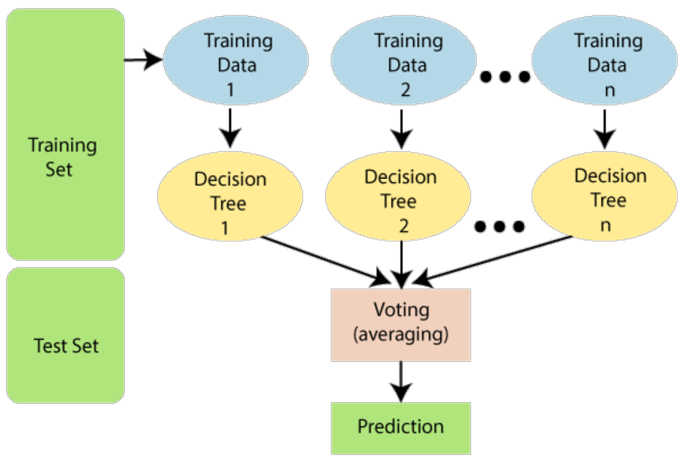

Figure 4.4: *A Random Forest Algorithm.*

Later we have utilised Random Forest to predict crop prices on our dataset. Following preprocessing, which involved one-hot encoding and handling of missing values, we divided the data into training and testing sets. We trained and tested the model using the Random Forest Regressor with 100 trees. We then used metrics like Mean Squared Error, Root Mean Squared Error, and R-squared to assess the model's performance. This method takes advantage of Random Forest's ensemble nature to capture intricate relationships and improve prediction accuracy.

4.2.3 XGBoost

XGBoost is a distributed gradient boosting library optimised for efficiency and scalability in machine learning model. It is an ensemble learning technique that generates a stronger prediction by combining the predictions of several weak models. XGBoost's effective handling of missing values is one of its main characteristics, enabling it to handle real-world data with missing values without requiring a lot of pre-processing.

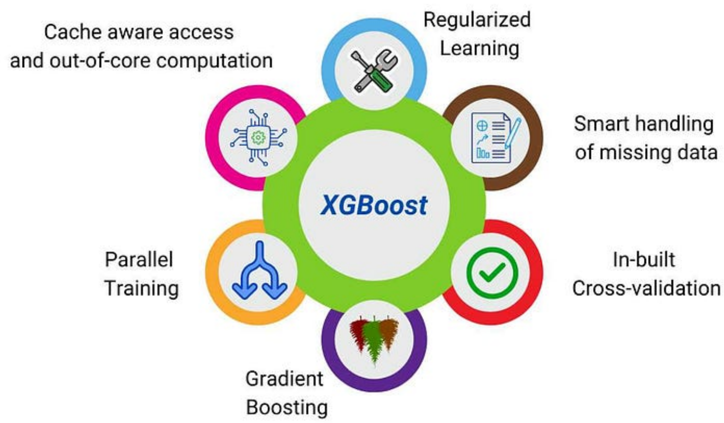

Figure 4.5: *The XGBoost Algorithm.*

Next we applied the XGBoost algorithm. The training data were then used to create and train an XGBoost Regressor model with 100 estimators and a random seed of 42. Metrics like Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R2) were used to evaluate the predictive performance of the model on the test data. XGBoost also performed well as it has that potential to handle missing values.

4.2.4 Ensemble Classifier

Following Random Forest, we used an ensemble classifier to improve predictive accuracy even further. The Decision Tree, Random Forest, and XGBoost models were all used in this ensemble classifier. Each model contributed to the final prediction, leveraging their distinct strengths and mitigating individual deficiencies, resulting in a more robust and accurate crop price prediction system. The ensemble approach is intended to capitalise on the various learning patterns of individual models, resulting in improved overall performance.
At first we have creates individual regression models using Decision Tree, Random Forest and multiple instances of XGBoost.

$$P_{ensemble} = \frac{1}{N}\sum_{i=1}^{N} P_i$$

where N is the number of models in the ensemble, and P_i is the prediction from the i-th model.

Then, it combines these models into a Voting Regressor ensemble. Finally, the ensemble model is trained on the training data, and its performance is evaluated on the test data using Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R2) metrics.

4.3 Regularization

Regularisation techniques are critical in machine learning because they prevent models from becoming overly complex and overfitting to training data. When a model learns the training data too well, it captures noise and specificities that do not generalise well to new, unseen data. Regularisation introduces a penalty term into the loss function of the model, discouraging overly complex models and encouraging generalisation. The Elastic Net regularisation technique was used in our crop price prediction project to improve the predictive power and robustness of our regression model. Here's a breakdown of how we implemented it.

4.3.1 Implementation

The main objective was to find the best hyperparameter alpha for Elastic Net regularisation. This procedure needed to strike the accurate balance between L1 and L2 regularisation, ensuring that our machine learning model achieves robust generalisation to new and previously unseen data. In order to represent various levels of regularisation strength, our method comprised defining a range of alpha values, ranging from 0.01 to 1.0 with increments of 0.01. An Elastic Net model was initialised for each alpha value, and 5-fold cross-validation was carried out with negative mean squared error (MSE) as the scoring metric. Selecting a negative mean square error (MSE) is consistent with the cross_val_score function's maximisation goal. Next, the alpha value that corresponded with the lowest cross-validation error was used to determine the ideal alpha. The MSE went from being negative to positive by eliminating the negative sign. The results were then printed for a thorough analysis, along with the best alpha, the associated MSE, and the R-squared (R2) score. The model's predictive performance and generalisation abilities were improved by fine-tuning the regularisation parameters.
Here is a visual representation of the trade-off between regularization strength (alpha) and model performance:

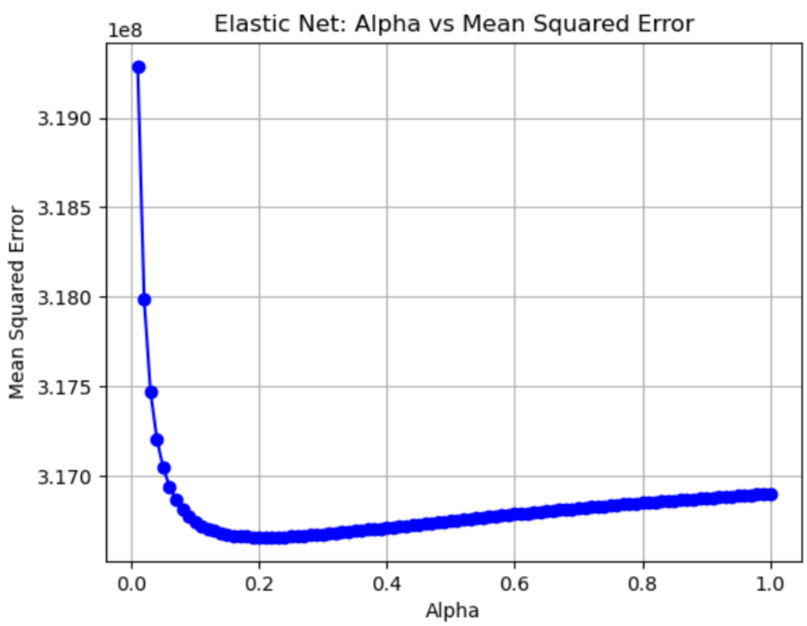

Figure 4.6: *Relationship between Alpha and Mean Squared Error in Elastic Net Regularization*

Chapter 5

Results and Analysis

We have used evaluation metrics, such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R2), to evaluate the performance of our model. Together, these measures offered a thorough evaluation of the model's accuracy, indicating the amount of variance the model could account for as well as the magnitude of prediction errors. MSE was used to show the average squared differences between the actual and predicted values, whereas RMSE, which calculated the square root of MSE, offered a more comprehensible measure. The amount of variance in the dependent variable that the model was able to account for was measured by R2, also known as the coefficient of determination.

1. **MSE(Mean Squared Error)**

   Model performance is better when the MSE value is lower. With the lowest MSE (12,909,159.57) in this aspect, the Random Forest model appears to be the most accurate in decreasing the prediction errors. The MSE values of the Decision Tree and XGBoost models are higher than those of Random Forest, indicating greater prediction error. Similar to Random Forest in terms of MSE, the Ensemble Classifier likewise does well in this parameter.

2. **RMSE (Root Mean Squared Error)** The square root of the mean square error, or RMSE, is a metric for predicting accuracy. Once more, smaller RMSE values are preferable. With the lowest RMSE values, Random Forest and Ensemble Classifier are thought to offer more accurate predictions. Though still reasonable, Decision Tree and XGBoost have somewhat larger RMSE values, indicating more substantial prediction mistakes.

3. **R2 Score (Coefficient of Determination)** The R2 Score quantifies the percentage of the dependent variable's variance that can be predicted based on the independent variables. Better predictive performance is indicated by higher numbers, which range from 0 to 1. Given their high R2 values of 0.96 in this instance, Random Forest and the Ensemble Classifier both show very good predictive power. Both Decision Tree and XGBoost exhibit commendable R2 scores: Decision Tree's score is 0.89, while XGBoost's is 0.95.

Based on these metrics, Random Forest and the Ensemble Classifier appear to be the top-performing models in this comparison, as they have the lowest MSE and RMSE values and the highest R2 scores. Decision Tree and XGBoost are also competitive but have slightly higher prediction errors and lower R2 scores in comparison.

Approach	MSE	RMSE	R2 Score
Decision Tree	40465820.64	6361.27	0.89
Random Forest	12909159.57	3592.93	0.96
XGBoost	15093640.81	3885.05	0.95
Ensemble Classifier	14872691.23	3856.51	0.96

Table 5.1: *Result Comparison*

Here is a comparison of MSE, RMSE and R2 Score for different algorithms we have utilized. From this we can see ensemble classifier which incorporates Decision Tree, Random Forest, and XGBoost performed better with 0.96 R2 Score.

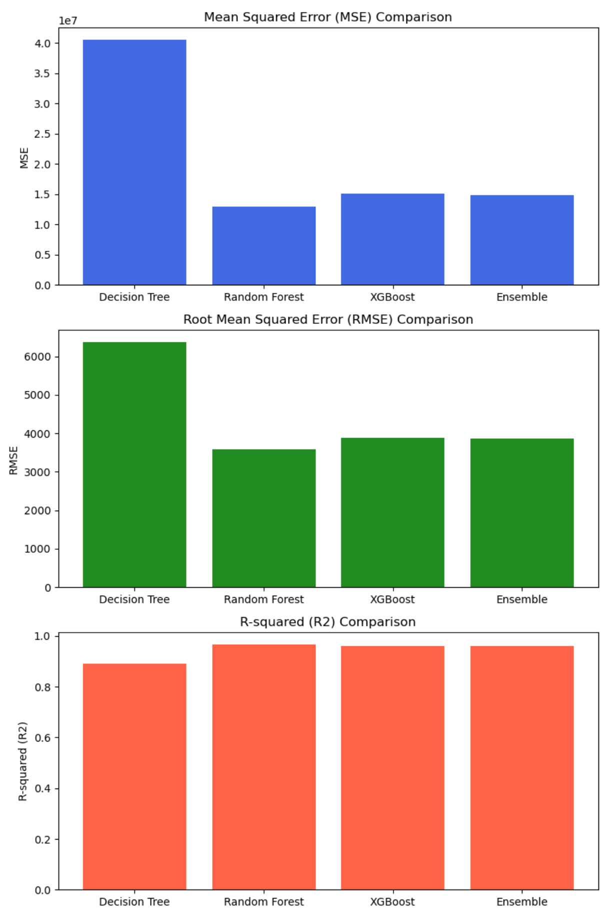

Figure 5.1: *MSE, RMSE, R2 metric comparison between these models*

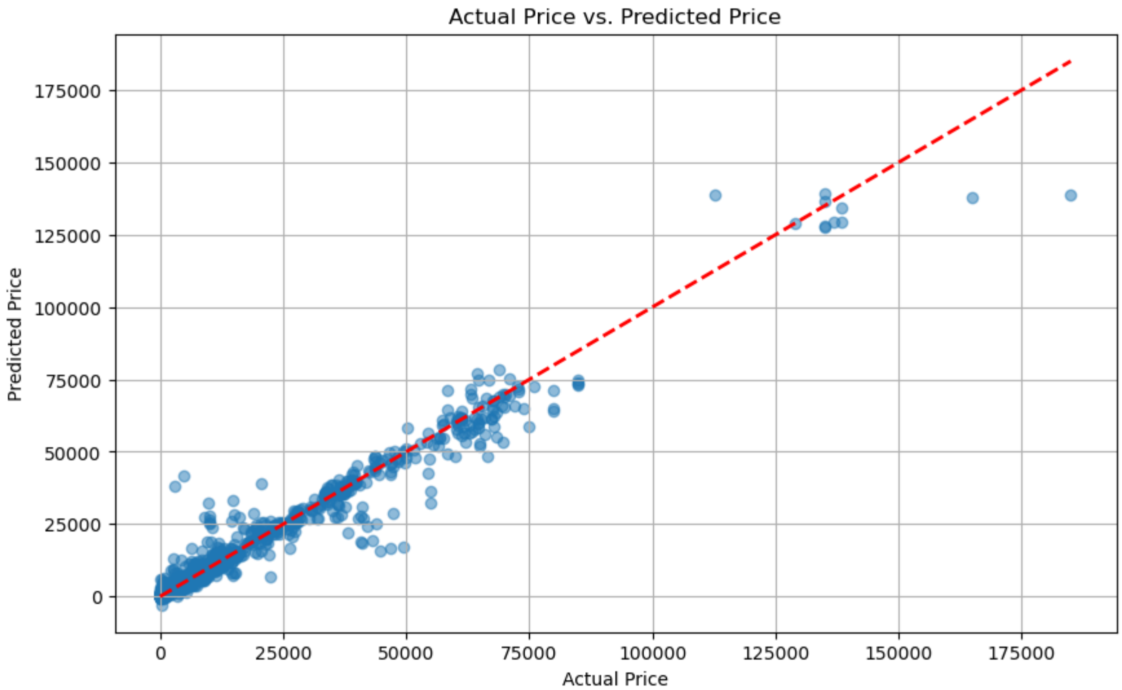

Figure 5.2: *Scatter Graph of Predicted Price vs Actual Price*

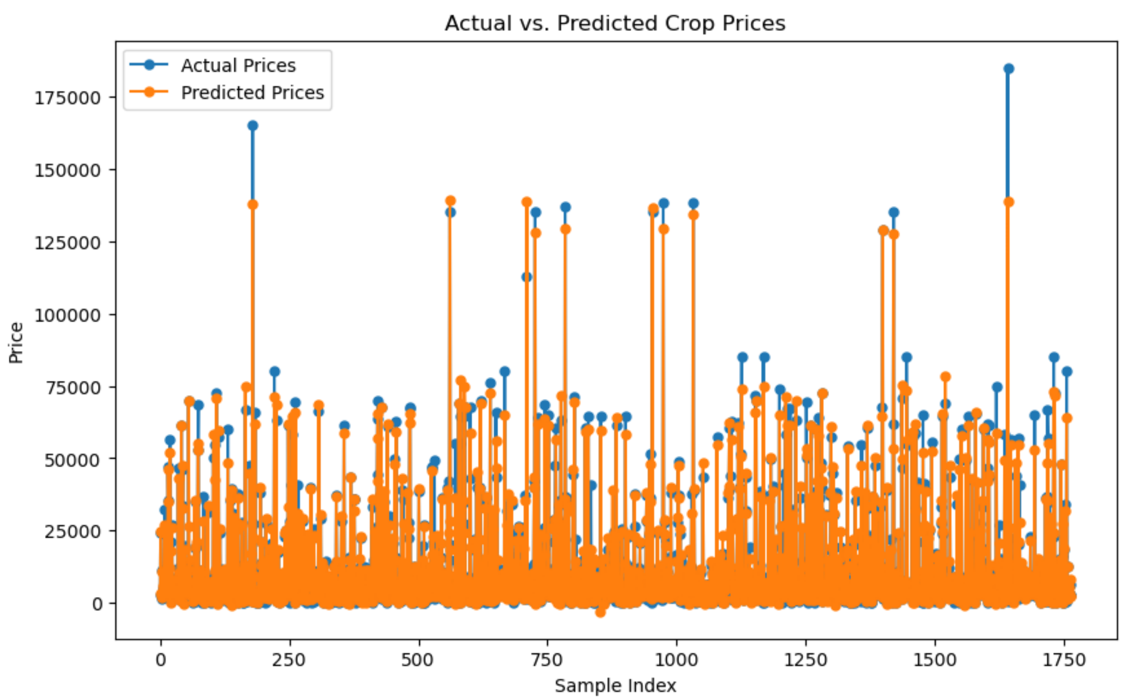

Figure 5.3: *Line Graph of Predicted Price vs Actual Price*

The scatter and line graph which representation the range of similarity between the actual and expected graphs here emphasise how well the model captured and developed the observed patterns, demonstrating its robust predictive ability. This relationship shows that crop prices were predicted with a high degree of accuracy using the selected features and machine learning algorithms.

The easy-to-use interface accepts user input and accurately predicts prices, demonstrating the model's accessibility and reliability.

```
for actual, predicted, diff in results:
    print(f"Actual Price: {actual:.2f}, Predicted Price: {predicted:.2f}")
```

```
Actual Price: 25250.00, Predicted Price: 25220.75
Actual Price: 2418.00, Predicted Price: 2388.12
Actual Price: 2316.00, Predicted Price: 2347.46
Actual Price: 10446.00, Predicted Price: 10477.47
Actual Price: 6383.00, Predicted Price: 6351.19
Actual Price: 5501.00, Predicted Price: 5468.62
Actual Price: 0.00, Predicted Price: 32.64
Actual Price: 2535.00, Predicted Price: 2501.50
Actual Price: 1927.00, Predicted Price: 1961.90
Actual Price: 2697.00, Predicted Price: 2731.97
Actual Price: 2937.00, Predicted Price: 2901.84
Actual Price: 3251.00, Predicted Price: 3215.73
Actual Price: 383.00, Predicted Price: 419.06
Actual Price: 3083.00, Predicted Price: 3119.07
Actual Price: 5521.00, Predicted Price: 5484.61
Actual Price: 4557.00, Predicted Price: 4520.42
Actual Price: 7332.00, Predicted Price: 7370.09
Actual Price: 3819.00, Predicted Price: 3780.63
Actual Price: 2425.00, Predicted Price: 2386.52
Actual Price: 5700.00, Predicted Price: 5739.65
```

Figure 5.4: *Actual price vs Predicted Price*

Chapter 6

Discussion

Our project's outcomes demonstrate a reliable crop price prediction system that was made possible by combining ensemble modelling. By addressing data difficulties transparently, improving forecast accuracy, and offering an easily navigated interface to stakeholders, this comprehensive approach ensures the system's efficiency in dynamic agricultural markets.

When contrasting our project with other noteworthy works, it distinguishes itself in a number of significant ways. Firstly, what distinguishes our method is the use of an ensemble classifier that combines XGBoost, Random Forest, and Decision Tree models. Compared to methods that rely on a single model, an ensemble strategy makes use of the strengths of individual models to generate a prediction system that is accurate and reliable.

Another noteable work of our project is dataset collection. Our dataset originates from the "Yearbook of Agricultural Statistics, Bangladesh", encompassing the years 2019, 2020, and 2021. This distinctive source enhances the uniqueness of our dataset, providing valuable and comprehensive agricultural statistics for a more robust analysis compared to other sources.

Our approach offersa holistic approach by combining ensemble modelling, optimisation via Elastic Net regularization and a user-friendly interface. When compared to other work in the field, these factors contribute to the novelty and effectiveness of our system. Our crop price prediction technology has the potential to accurately predict crop prices, providing farmers and other agricultural stakeholders with crucial assistance.

Chapter 7

Limitations

Data Accessibility and Quality: The accuracy of predictions may be impacted by limited access to high-quality historical data for model training.

Market Manipulation: Crop prices can be affected by speculative activity or possible market manipulation, which the model might not take into account.

Features with a Limited Scope: Unforeseen variables that affect crop prices but are not accounted for in the model may result in inaccurate predictions.

Stationarity Assumption: Assuming that historical patterns will repeat themselves in the future (stationarity) may not be valid, particularly in dynamic agricultural markets.

External Factors: Crop prices can be greatly impacted by outside variables that the model may not fully account for, such as natural disasters, or political events.

Dependence on previous Data: The model is strongly dependent on previous data, and changes in market dynamics may make past patterns less predictive of future trends.

Chapter 8

Ethical Concerns

1. There could be an ethical issue if we don't take consent from the farmers and others for our dataset but we have gathered our data from the "Year Book of Agricultural Statistics of Bangladesh" and we have asked permission before using their information and we make sure to keep it private and safe and we have clarified who owns this data.

2. The absence of a clear dataset license raises ethical concerns regarding data ownership but in our research, We have gathered data from the "Year Book of Agricultural Statistics of Bangladesh" so the data is reliable and there is assurance that proper licensing has been secured for the dataset.

3. There can be another ethical issue of data privacy if the personal information of farmers and stakeholders is disclosed but in our research, we are committed to safeguarding data privacy by anonymizing or aggregating information protecting the identities of individual farmers and districts.

4. Another issue can be bias and fairness of data that could impact the fairness of prediction acknowledging the potential biases inherent in historical data, we actively assess and mitigate biases to ensure fairness in predictions.

5. There could be an environmental impact such as energy consumption for computation, data storage, and transmission, as well as considerations related to electronic waste and the sustainability of long-term model maintenance. We care about the impact our project might have on people and the environment, and we're doing our best to make sure it's helpful and doesn't cause any problems.

6. Another ethical concern may arise in software from inadequate user privacy protection, biased algorithms, or lack of transparency in decision-making processes. In our research work, we have ensured a transparent decision-making process and we have tried to mitigate algorithmic biases.

Chapter 9

Conclusion

In conclusion, our crop price prediction system entailed an in-depth analysis of various machine learning algorithms. We overcame issues with data accessibility and quality by utilizing a dataset from the "Yearbook of Agricultural Statistics, Bangladesh" that covered the years 2019, 2020, and 2021. Extensive data were methodically collected from diverse districts across Bangladesh, providing a comprehensive picture of agricultural dynamics. Crop pricing, meteorological parameters (rainfall, temperature, humidity), and pesticide use were all taken into account. Our work used an ensemble method, combining the advantages of Random Forest, XGBoost, and Decision Tree models to produce a reliable prediction system. We improved our models by using regularization techniques like Elastic Net, which highlights the significance of hyperparameter tuning. A user-friendly interface was made possible by the integration of Streamlit for web frontend, allowing users to enter commodity details and get accurate price predictions. Our work recognizes limitations related to market dynamics, external factors, and the assumption of data stationarity, Our distinct data collection method and the use of ensemble classifiers are highlighted by the comparison with previous works, which represents a noteworthy contribution to the field. All things considered, our research provides insightful information about predicting crop prices and paves the way for further developments and advancements in agricultural estimations.

9.1 Future Scope:

Data Augmentation: Apply data augmentation techniques to expand the dataset's size and diversity.

Including Real-Time Data: Incorporate real-time data streams, including social media signals, satellite imagery, and weather updates, to improve the model's potential to adapt to dynamic changes in market and agricultural conditions.

Adaptive Feature Engineering: Develop models with dynamic feature engineering so the system can adjust to changing market dynamics.

Collaboration with Specialists: Work together with agricultural specialists to gain a deeper comprehension of the complexities of the industry.

Bibliography

[1] Yearbook of agricultural statistics, bangladesh.

[2] World Bank Group. Bangladesh: Growing the economy through advances in agriculture. World Bank Group, Oct 2016.

[3] Bangladesh economic review 2022, agriculture. Ministry of Finance-Government of the Peoples Republic of Bangladesh.

[4] Jahidul Islam11 August and Jahidul Islam. Middlemen eat into farmers' pie. Aug 2020.

[5] B Chaitra and K Meena. Forecasting crop price using various approaches of machine learning. In *2023 International Conference on Innovations in Engineering and Technology (ICIET)*, pages 1–5. IEEE, 2023.

[6] G Thapaswini and M Gunasekaran. A methodology for crop price prediction using machine learning. In *2022 IEEE 2nd International Conference on Mobile Networks and Wireless Communications (ICMNWC)*, pages 1–7. IEEE, 2022.

[7] S Sajithabanu, A Ponmalar, A Gnana Soundari, NV Reshma, K Sunraja, and R Sindhumathi. Enhanced crop price prediction & forecasting system. In *2022 International Conference on Computer, Power and Communications (ICCPC)*, pages 580–585. IEEE, 2022.

[8] Md Ishak, Md Shahidur Rahaman, and Tahasin Mahmud. Farmeasy: An intelligent platform to empower crops prediction and crops marketing. In *2021 13th International Conference on Information Communication Technology and System (ICTS)*, pages 224–229, 2021.

[9] Ramesh Medar, Vijay S Rajpurohit, and Shweta Shweta. Crop yield prediction using machine learning techniques. In *2019 IEEE 5th international conference for convergence in technology (I2CT)*, pages 1–5. IEEE, 2019.

[10] Ishita Ghutake, Ritesh Verma, Rohit Chaudhari, and Vidhate Amarsinh. An intelligent crop price prediction using suitable machine learning algorithm. In *ITM web of conferences*, volume 40, page 03040. EDP Sciences, 2021.

[11] Dhanasekaran K, Ramprasath M, Sathiyamoorthi V, Poornima N, and Irrai Anbu Jayaraj. Meta-learning based adaptive crop price prediction for agriculture application. In *2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pages 396–402, 2021.

[12] Ye Lu, Li Yuping, Liang Weihong, Song Qidao, Liu Yanqun, and Qin Xiaoli. Vegetable price prediction based on pso-bp neural network. In *2015 8th International Conference on Intelligent Computation Technology and Automation (ICICTA)*, pages 1093–1096, 2015.

[13] Pundru Chandra Shaker Reddy, G. Suryanarayana, LNC Prakash K, and Sucharitha Yadala. Data analytics in farming: Rice price prediction in andhra pradesh. In *2022 5th International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT)*, pages 1–5, 2022.

[14] K. Priyadharshini, R. Prabavathi, V. Brindha Devi, P. Subha, S.Mohana Saranya, and K. Kiruthika. An enhanced approach for crop yield prediction system using linear support vector machine model. In *2022 International Conference on Communication, Computing and Internet of Things (IC3IoT)*, pages 1–5, 2022.

[15] Kajal P Parmar and Tejas Bhatt. Crop yield prediction based on feature selection and machine learners: A review. In *2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS)*, pages 354–358, 2022.