

Programming and Data Structures Project

-B.tech First Year

Crop Price Prediction

using c language

-by Pranav ah P.dhanush



INDIAN INSTITUTE OF TECHNOLOGY
BHUBANESHWAR

Project Report

i. Motivation: Why did you choose the project?

The motivation behind choosing this project is from my grandparents , we had a farm in our native where they grew many crops including rice ,wheat,, millets, pulses etc .

Agriculture plays an important role in India's economy, and the importance of predicting crop prices for farmers, traders, and policymakers is immense. By analyzing historical crop data and using regression techniques, we aimed to provide insights into how factors like temperature, rainfall, and demand influence the prices of major crops. This project offers a practical solution for price forecasting, which can be useful for strategic decision-making in agriculture.

ii. Important Highlights of the Project

- **Data-Driven Price Prediction:** We used historical crop data for different states of India to predict the future prices of crops such as rice, wheat, corn, pulses, and millets.
- **Multiple Linear Regression:** Implemented a multivariate linear regression model considering temperature, rainfall, and population-driven demand as input variables.
- **Dynamic Crop and State Selection:** The model allows users to input their desired crop and state to obtain a price prediction based on current conditions.
- **Scalable Approach:** The project is designed to be scalable by using CSV files for data storage, which allows for easy integration of additional data points in the future.

.

iii. What did you learn from this exercise? Concepts that you explored and implemented outside of your comfort zone.

Through this project, I gained a deeper understanding of:

- **Data Parsing and Handling:** Reading large CSV files and efficiently handling missing or erroneous data was a challenge. I had to ensure that all data points were correctly mapped to the relevant variable.
- **Multiple Linear Regression:** I explored how regression works for multiple variables and how each factor (temperature, rainfall, demand) affects the output (price). This was new to me and involved implementing the mathematical foundation of regression.
- **Demand Estimation:** The use of population as a proxy for demand was a concept I hadn't previously worked with, and it required careful calculation and validation.
- **Prediction Modeling:** Learning how to structure the prediction logic in C and making sure the model's output made sense in real-world scenarios was an interesting learning curve.
- **Implementing new features:** Learning how to read data from a file, how to implement use of files in programming, structures usage.

iv. Areas of Improvement: What things you could have improved on provided more time/experience?

- **Data Accuracy and Cleaning:** Given more time, I would focus on improving the data cleaning process to handle missing or corrupted values more efficiently. I would also explore better ways to interpolate data gaps.
- **Model Sophistication:** The current model is limited to linear regression. Given more time, I would experiment with more advanced machine learning models, such as Random Forest or Neural Networks, to improve the accuracy of price predictions. Or we could have also used more adaptive regression techniques polynomial regression , Non-linear regression etc.
- **Incorporating More Variables:** I would incorporate more variables such as government policies, subsidies, international market trends, availability of land, cheap labour, expense reduction and crop health indices (NDVI) to increase the model's robustness.
- **User Interface:** Enhancing the user interface to make it more interactive and user-friendly could make this project accessible to a wider audience, particularly farmers and agricultural planners.
- **Larger scale:** The above code can be implemented for larger number of crops including fruits, vegetables etc could be for livestock also .
- **Improve accuracy:** We could include cities instead or states for more accurate results, or even districts.
- **Larger data:** We can increase the number of data points directly uploading a bigger csv file having more data for previous years.

v. Future Scope: How do you foresee this project being useful in the future?

This project has the potential to evolve into a more comprehensive tool for agricultural decision-making. By improving the model with more sophisticated algorithms and incorporating real-time data streams, it could help farmers predict optimal selling times and pricing strategies. It could also be useful for governmental and non-governmental organizations in formulating policies related to agricultural subsidies, procurement, and market interventions.

vi. Contribution of Each Member to the Project

- **Pranav:** Responsible for implementing the multiple linear regression algorithm, data handling, formatting csv file data and prediction logic. Managed the testing and debugging of the system.
- **Dhanush:** Proposed project idea, assisting with data collection and organisation, supporting debugging and testing, conducting data prediction and analysis.

Vii CONSTRAINTS - The above code is applicable for 10 states and 4 crops.

• STATES-

Andhra Pradesh

Bihar

Gujarat

Karnataka

Madhya Pradesh

Maharashtra

Punjab

Rajasthan

Tamil nadu

Uttar Pradesh

• **CROPS -**

Rice

Wheat

Corn

Pulses

Millet