Predicting Future Farm Product Prices

James Bootsma, Aditya Manickam Arumugam, Krishang Karir, Abdul Aziz Essam, Mohd Daiyaan Nasir

**Description of Applied Problem**

Accurate predictions of farm product prices are crucial for the agricultural sector. Therefore, this project aims to develop a predictive model using historical climate[1, 2] and economic[3] data to forecast future prices. Empowering stakeholders to make informed decisions, optimize productivity, and enhance resilience against economic fluctuations in the agricultural sector.

**Description of Available Data**

1. **Farm Product Price Data -** Monthly prices of various farm **products** across Canada.[4]
2. **Historical Climate Data -** Monthly climate statistics including temperature, precipitation, and snowfall by weather station.[5]
3. **Agricultural Gross Domestic Product -** GDP of crop and animal production.[6, 7]
4. **Oil Prices -** Monthly average oil prices across Canadian cities.[8]

**Plan for Analysis and Visualization**

1. **Data Preprocessing:**
   * The data sources climate and oil prices from multiple locations in each province and territory. We aggregated this data using the mean by province and territory.
   * Monthly GDP by province is calculated by distributing the monthly national agricultural GDP based on each province's annual agricultural GDP share.
2. **Feature Engineering:**
   * Develop features such as seasonal averages, extreme weather events, and lagged climate variables to capture their impact on farm product prices.
   * Calculate additional economic features (e.g., GDP growth rate, oil price changes) to incorporate economic context.
3. **Exploratory Data Analysis (EDA):**
   * Visualize the relationships between climate variables, economic indicators, and farm product prices using scatter plots, heat maps, and time series plots.
4. **Model Development:**
   * **Machine Learning Algorithms:**
     + Explore regression (e.g., Linear Regression, Random Forest Regression[9], SVM) and time series models[10] (e.g., ARIMA, LSTM) to predict future prices.
     + Use cross-validation to tune model parameters and assess performance.
   * **Model Evaluation:**
     + Evaluate performance using metrics such as MAE, RMSE, and R-squared.
     + Compare a model trained on data from all products with a model trained on a single product to determine which approach yields better predictions.
5. **Visualization and Reporting:**
   * Create dashboards for the user to track how different products are being predicted and how they react to various factors.

**References:**

1. Schlenker, W., & Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to US crop yields under climate change. Proceedings of the National Academy of Sciences, 106(37), 15594-15598.
2. Lobell, D. B., & Field, C. B. (2007). Global scale climate–crop yield relationships and the impacts of recent warming. Environmental Research Letters, 2(1), 014002.
3. Fernando Avalos, (2014). Do oil prices drive food prices? The tale of a structural break, Journal of International Money and Finance, Volume 42, Pages 253-271, ISSN 0261-5606.
4. Statistics Canada. [Table 32-10-0077-01  Farm product prices, crops and livestock](https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3210007701)
5. Government of Canada. (2024, June 27). [Monthly Climate Summaries. Climate.](https://climate.weather.gc.ca/prods_servs/cdn_climate_summary_e.html)
6. Statistics Canada. [Table 36-10-0434-01  Gross domestic product (GDP) at basic prices, by industry, monthly (x 1,000,000)](https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3610043401)
7. Statistics Canada. [Table 36-10-0402-02  Gross domestic product (GDP) at basic prices, by industry, provinces and territories, growth rates (x 1,000,000)](https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3610040202)
8. Statistics Canada. [Table 18-10-0001-01  Monthly average retail prices for gasoline and fuel oil, by geography](https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1810000101)
9. Zheng, J., Xin, D., Cheng, Q., Tian, M., & Yang, L. (2024). The random forest model for analyzing and forecasting the US stock market under the background of Smart Finance. Atlantis Highlights in Computer Sciences, 82–90.
10. Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: principles and practice.

## **products**

(might select just a few)

* Barley
* Barley for animal feed
* Barley for malt and other human consumption
* Calves for feeding
* Calves for slaughter
* Canadian Wheat Board, barley excluding payments
* Canadian Wheat Board, barley including payments
* Canadian Wheat Board, durum excluding payments
* Canadian Wheat Board, durum including payments
* Canadian Wheat Board, selected barley excluding payments
* Canadian Wheat Board, selected barley including payments
* Canadian Wheat Board, wheat excluding payments
* Canadian Wheat Board, wheat including payments
* Canary seeds
* Canola (including rapeseed)
* Cattle for feeding
* Cattle for slaughter
* Chickens for meat
* Corn for grain
* Cows for slaughter
* Dry peas
* Durum wheat
* Eggs in shell
* Flaxseed
* Fresh potatoes
* Fresh potatoes for processing
* Fresh potatoes for seed
* Fresh potatoes for table consumption
* Heifers for feeding
* Heifers for slaughter
* Hogs
* Lambs
* Lentils
* Non-board wheat (except durum wheat)
* Oats
* Ontario wheat excluding payments
* Ontario wheat including payments
* Rye
* Soybeans
* Steers for feeding
* Steers for slaughter
* Turkeys for meat
* Unprocessed milk from bovine
* Wheat (except durum wheat)
* Wheat (except durum wheat), milling
* Wheat (except durum wheat), other