Article

Not peer-reviewed version

Using probabilistic machine learning methods to improve beef cattle price modeling and promote beef production efficiency and sustainability in Canada

Elham Rahmani * , Mohammad Khatami , Emma Stephens *

Posted Date: 14 December 2023

doi: 10.20944/preprints202312.1090.v1

Keywords: Machine Learning, Probabilistic Modeling, Multivariate and Univariate Modeling, Support Vector Regression, Random Forest, Adaboost, ARIMA, SARIMA, SARIMAX, Canadian Cattle Price Modeling

Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 December 2023 doi:10.20944/preprints202312.1090.v1

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Using Probabilistic Machine Learning Methods to Improve Beef Cattle Price Modeling and Promote Beef Production Efficiency and Sustainability in Canada

Elham Rahmani [1],*, Mohammad Khatami [2] and Emma Stephens [1],*

[1] Agriculture and Agri-Food Canada, Lethbridge Research and Development Centre, Lethbridge, Alberta, Canada; elham.rahmani@agr.gc.ca

[2] B-IT and Department of Computer Science, University of Bonn, Bonn, Germany; mdkhatami@gmail.com

* Correspondence: Autours: emma.stephens@agr.gc.ca & elham.rahmani@agr.gc.ca

Abstract: Accurate agricultural commodity price models enable efficient allocation of limited natural resources, leading to improved sustainability in agriculture. Because of climate change, price volatility and uncertainty for the sector is expected to increase in the future, increasing the need for improved price modeling. With the emergence of machine learning (ML) algorithms, novel tools are now available to enhance the modeling of agricultural commodity prices. This research uses both univariate and multivariate ML techniques to perform probabilistic price prediction modelling for the Canadian beef industry, taking into account beef production, commodity market and international trade features to enhance accuracy. We apply multivariate algorithms by using support vector regression (SVR), random forest (RF), and adaboost (AB) and univariate algorithms by using autoregressive integrated moving average (ARIMA), seasonal ARIMA (SARIMA), and the seasonal autoregressive integrated moving average with exogenous factors (SARIMAX). We apply these models to monthly fed steer price data between January 2005 and September 2023 and compare predicted prices with observed prices using several validation metrics. The outcomes indicate that both random forest (RF) and adaboost (AB) in multivariate modeling, show superior performance in accurately predicting Alberta fed steer prices in comparison to other algorithms. To better account for the variance of the best model performance, we subsequently adopted a probabilistic approach, by considering uncertainty to our best selected ML model. The beef industry can use these improved price models to minimize resource waste and inefficiency in the sector and improve the long-term sustainability prospects for beef producers in Canada.

Keywords: Machine Learning; Probabilistic Modeling; Multivariate and Univariate Modeling; Support Vector Regression; Random Forest; Adaboost; ARIMA; SARIMA; SARIMAX; Canadian Cattle Price Modeling

1. Introduction

Agricultural commodity markets are often uncertain and unpredictable and can be affected by a wide variety of factors, including but not limited to: agricultural input prices and production conditions; fuel and other commodity price swings; agriculture industry financial factors; weather, natural disasters, and climate change; the global economy; and political shocks. Consequently, reliable and timely agricultural commodity price modeling is critical in ensuring the sustainability and economic viability of the agricultural sector by providing better information on commodity price behavior. The increasing availability of ever larger and more comprehensive sets of agricultural data, as well as the consistent need for accurate commodity price models, necessitates the development of robust and efficient analysis techniques that can be used to improve our understanding of commodity prices from current observations.

Agricultural commodity price modeling has a long history in agricultural economics, and many different methods have been applied. These have included using time series econometric models, tracking futures prices, and expert opinion and qualitative assessments. With the advent of machine learning, this offers new techniques to analyze and use for commodity price modeling. Identifying

© 2023 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 December 2023 doi:10.20944/preprints202312.1090.v1

the appropriate machine learning methods for agriculture price analysis has gotten less attention than the application of ML to other non-agricultural sectors [1,2] and thus more research on machine learning strategies appropriate for agricultural price analysis is required [3].

To contribute to this research, we make use of the long-term historical cattle price data available from the CanFax research service. CanFax is an established market research firm that specializes in analysis of the Canadian beef sector (www.canfax.ca/Research.aspx). We structure fed cattle price prediction as a machine learning problem using key price and production data for the Canadian beef sector from CanFax and several other sources, and assess the ability of widely used machine learning algorithms to predict observed fed cattle prices using information from important correlated data. We also generate price predictions with realistic variance by modeling the probability distribution of the residual and considering it as an error term. We assume that the actual prices and residuals are realizations of a normally distributed random variable, characterized by an average of zero and a variance derived from the differences between the actual and predicted data. This allows us to consider uncertainty in our ML modeling [4].

Machine learning methods are well suited to discovering complex relationships and hidden patterns across multivariate datasets. This flexibility and pattern recognition capacity minimizes the errors that can arise from incorrect application of structural models of agricultural commodity price processes. The following is a summary of our main contributions in this study:

- With the long-term historical data available on cattle prices, agricultural commodity market, economical indices, and international trade features, we structure price prediction as a machine learning problem, which can be more accurate, consistent, and efficient than traditional time series statistical methods.
- Three multivariate machine learning algorithms; support vector regression, random forest regression, and adaboost regressor models and three univariate time series algorithms; ARIMA, SARIMA, and SARIMAX were applied to long-term historical fed cattle price datasets (2005 to 2023) in Alberta. We assess the performance of these algorithms against observed fed cattle prices and identify the best machine learning algorithm for fed cattle price prediction.
- The multivariate machine learning approach offers a feasible alternative to structural multivariate autoregressive modeling and can efficiently combine fed beef price and exchange rate dynamics, minimizing modeling errors and data requirements

The rest of the paper is organized as follows: Section 2 reviews existing research on agricultural commodity price analysis, machine learning applications in the agricultural sector and research on commodity price analysis and sustainability outcomes. Section 3 discusses the materials and methods used in this study, including the study domain and data descriptions, variables used, data preprocessing and partitioning, an introduction to each of the algorithms, as well as model validation. Section 4 discusses the findings of both multivariate and univariate models, as well as feature engineering and hyperparameter tuning results. In this section we apply probabilistic modeling approach to the best selected ML model. Section 5 concludes.

2. Background Literature

Agricultural commodity price modeling has been performed for many decades, employing a wide variety of methods, from expert opinion outlooks and time series analysis [5–7] to more recent forays into machine learning approaches [2,8].

Livestock price modeling has been an important segment of agricultural price analysis work, particularly in North America, and prior research has demonstrated that useful models depend on assessing a complex set of factors including animal health and production dynamics, seasonal, cyclical and spatial patterns [9], feed grain production conditions [7,10]and even finance and macroeconomic conditions for regions where livestock exports and imports are important [6].

A variety of methods have been employed, from univariate time series models to multivariate time series and structured multiple equation models of both livestock prices as well as the prices of important related inputs, like feeds. An early example of univariate modeling by Oliveira, *et al.* [11] uses the ARIMA Box-Jenkins approach to produce short-run forecasts of live cattle prices and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 December 2023 doi:10.20944/preprints202312.1090.v1

compares these to futures prices and finds that ARIMA models do well at short horizons. Zapata and Garcia [12] compare multivariate vector auto-regression and error correction (VAR, VEC) approaches to predict cattle slaughter prices in the United States based on monthly data from lagged prices as well as feeder prices and per capita consumer incomes. Several other statistical approaches are outlined in detail in Linnell [13].

Beyond statistical approaches, other analysis is produced for livestock markets worldwide based on expert opinions from knowledgeable stakeholders and either public or private advisers in the Livestock sector. Farm Credit Canada provides outlook reports, and the CanFax Research Service produces a weekly expert report for subscribers, combining quantitative and qualitative assessments of the livestock markets in Canada and the U.S. (canfax.ca). The United States Department of Agriculture, as well as several individual U.S. state agricultural departments, produce a wide variety of publicly available reports on livestock market conditions, which is important for Canadian producers given the high level of integration between Canadian and U.S. livestock markets. Many researchers also argue that ensemble analysis that combine and average the results from several different methodologies are better for assessing livestock market performance [5,7,13].

Evidence in the literature of prior use of machine learning to analyze cattle prices is limited. Kohzadi, *et al.* [14] employed a walk-forward or sliding window technique on commodity price data from 1970 through 1990 to compare artificial neural networks to time series models for modeling commodities prices of cattle and wheat. They found that the ARIMA model did not perform as well as the neural network models. The ARIMA model was most comparable for wheat, but the neural network models were able to catch a substantial number of turning points for both wheat and cattle. Beyond these references, there appears to be minimal exploration in the academic literature of machine learning for livestock price analysis, particularly in Canada. However, there appears to be increasing uptake of machine learning models in the private sector to provide commodity price modeling services more broadly, but within this, agricultural commodity prices again do not appear to be a major focus. More literature review on the various research and applications of these techniques for the agricultural sector is summarized in Table (1).

Table 1. Summary of application of ML to agriculture.

Autor/s	Year	Study Domain	Considered Parameter/s	Machine Learning Technique
Jeong et al.	2022	South Korea	Rice yield	five different structures of deep learning [15]
Sharma et al.	2021	Review paper	Precision agriculture	a comprehensive review [16]
Liu et al.	2021	Review paper	Precision agriculture	a systematic literature review [17]
Maroli et al.	2021	Review paper	sustainability in agricultural sector	a comprehensive review [18]
Meshram et al.	2021	Review paper	Pre-harvesting, harvesting and post-harvesting parameters	a comprehensive review on all ML techniques [19]
Chen et al.	2021	Malaysia	Agriculture commodity price	ARIMA, support vector regression (SVR), Prophet, extreme gradient boosting (XGBoost), long short-term memory (LSTM) [20]
Tian et al.	2021	Shaanxi, China	Wheat yield	long short- term memory (LSTM), back propagation neural network (BPNN), support vector machine (SVM) [21]
Divisekara et al.	2020	Canada, Saskatchewan	Forecasting the red lentils commodity market price	SARIMA models [22]

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 December 2023 doi:10.20944/preprints202312.1090.v1

Sharma et al.	2020	Review paper	Sustainable agriculture supply chain performance	a systematic literature review [23]
Kamir et al.	2020	Australia	Wheat yield	random forest (RF), cubist (CU), XGBoost (XGB), multi-layer perceptron (MLP), support vector regression linear (SVMl), support vector regression radial (SVMr), Gaussian process regression (GPR), k-nearest neighbor (kNN), multivariate adaptive regression (MARS) [24]
Yamaç & Todorovic	2020	Bari, Southern Italy	Daily potato crop evapotranspiration	k-nearest neighbour (kNN), artificial neural networks (ANN), adaptive boosting (AdaBoost) [25]
Van Klompenburg et al.	2020	Review paper	Crop yield prediction	a systematic literature review on artificial neural network (ANN) methods [26]
Vidyarthi et al.	2020	Kettleman, California	Size and mass of pistachio kernels	random forest (RF) [27]
Cai et al.	2019	Australia	Wheat yield	LASSO, support vector machine (SVM), random forest (RF), neural network (NN) [28]
Kouadio et al.	2018	Southern Vietnam	Robusta coffee yield	extreme learning machine (ELM), multiple linear regression (MLR), random forests (RF) [29]
Prajapati & Kathiriya	2016	Vadodara in western India	Soil health card	K-nearest neighbor (kNN) classification using nine different similarity measures [30]

2.1. Agricultural Commodity Price information and sustainability

Statistics Canada and organizations such as Canfax routinely provide livestock market data to producers and agricultural stakeholders to help guide decision making about beef production, marketing and trade. Availability of accurate agricultural price data has long been regarded as a key factor in promoting a successful and productive agricultural sector [31]. Accurate agricultural price information monitoring has been associated with securing global objectives on food security and environmental sustainability [32]. Thus using new methodological tools such as ML to analyze agricultural commodity prices can supplement these broader objectives to increase global access to accurate agricultural commodity market information to improve both production efficiency and sustainability.

3. Materials and Methods

3.1. Study Domain

The study domain of this research is Alberta, a province in the west of the Canadian Prairies. Since the late 1950s, the cattle industry rose to prominence and is a key component of the province’s agricultural sector . Canada contributes significantly to the trade and consumption of beef around the world, and approximately 11.5 million cattle exist in Canada, including 9.5 million beef and 2.0 million dairy cattle. Canada is the world's 12th largest beef producer, with 1.50 million tones (2% of world totals) of beef produced in 2020 [33]. Canada exports 47.4% of its beef, accounting for 4.8% of worldwide exports in 2020 and ranks seventh among beef exporters. Western Canada has 3.2 million beef cows and 79% of Canada's fed cattle is finished for slaughter [33]. Alberta has the largest average herd size (255 head) with 1.5 million (41%) beef cows in total, followed by Saskatchewan (1.1 million

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 December 2023 doi:10.20944/preprints202312.1090.v1

beef cows, 191 head/producer) (31%), and Manitoba (412 thousand beef cows, 167 head/producer) (11%) [33].

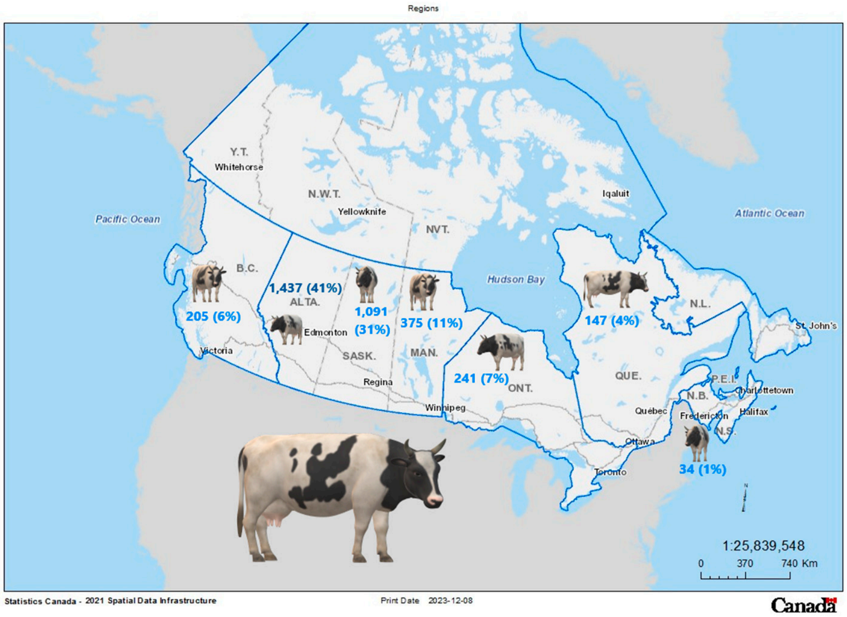


Figure 1. Canada's beef cows in provinces (all inventories in 1,000 head), as of January 2021 reported by StatCan[1] .

3.2. Data Acquisition & Description & Exploration

This study uses historical cattle price data gathered from CanFax[2], a market research firm which is a much-relied upon source of cattle market information in Canada. CanFax Research Services (CRS) delivers comprehensive statistical and market information on domestic and worldwide beef trends to the Canadian beef sector. The dataset includes regularly updated monthly cattle prices (Cdn$/cwt) starting in January 2005 to September 2023, for fed steer cattle in Alberta, along with several other cattle classes. In order to perform a multivariate machine learning analysis for predicting cattle prices, we constructed a variable matrix comprising several key data series known to be related to cattle prices. For our analysis, we include the consumer price index, which is a measure of overall price inflation, for all items in Canada[3], the monthly average Alberta natural gas price[4] ($CAD/gigajoule (GJ)) which is strongly related to agricultural production costs, the Canadian-US dollar exchange rate[5], and Alberta barely prices[6] ($CAD/tonne), as barley is the main feed grain used in Alberta beef production. Table 2 summarizes annual averaged values for these variables. Figures 2 and 3 also visualize the time series trends of fed steer prices and these related variables that we used for predicting cattle prices from January 2005 to September 2023. The natural gas price and exchange rate display the highest volatility over time. Meanwhile, the Fed steer price, barley price, and Canadian consumer price index show a noticeable steep increase in trend starting from 2020, coinciding with the onset of the COVID-19 pandemic outbreak.

[1] https://canadabeef.ca/wp-content/uploads/2021/09/Canada-Beef-Fast-Fact-Sheet-2021.pdf

[2] Research (canfax.ca)

[3] https://fred.stlouisfed.org/series/CPALCY01CAM661N

[4] https://economicdashboard.alberta.ca/dashboard/natural-gas-price#

[5] https://fred.stlouisfed.org/series/EXCAUS

[6] https://economicdashboard.alberta.ca/dashboard/wheat-price#Product:Barley

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 December 2023 doi:10.20944/preprints202312.1090.v1

Table 2. Annual average price values for key data variables (* 2023 includes Jan. to Sep.).

Year	Fed Steer Price ($/cwt)	Nat Gas Price ($/GJ)	Barley Price ($/tonne)	Exchange (CAD/US)	Canadian Consumer Price Index
2005	85.60	7.87	91.79	1.21	85.44
2006	86.90	6.22	101.81	1.13	86.93
2007	88.51	6.05	156.99	1.07	88.30
2008	90.07	7.47	199.31	1.07	90.15
2009	85.72	3.65	152.77	1.14	90.85
2010	89.08	3.57	138.07	1.03	91.93
2011	106.47	3.28	174.58	0.99	94.51
2012	112.32	2.14	214.24	1.00	96.39
2013	119.26	2.83	239.08	1.03	97.17
2014	156.51	4.00	173.54	1.10	98.65
2015	184.16	2.42	218.80	1.28	100.00
2016	153.75	1.83	215.96	1.33	100.81
2017	154.93	2.02	193.91	1.30	101.96
2018	153.68	1.29	213.36	1.30	103.68
2019	149.87	1.40	227.00	1.33	106.03
2020	138.94	1.90	215.26	1.34	107.04
2021	155.88	3.10	279.61	1.25	111.04
2022	173.08	4.78	380.14	1.30	118.44
2023*	222.36	2.57	386.30	1.35	124.01

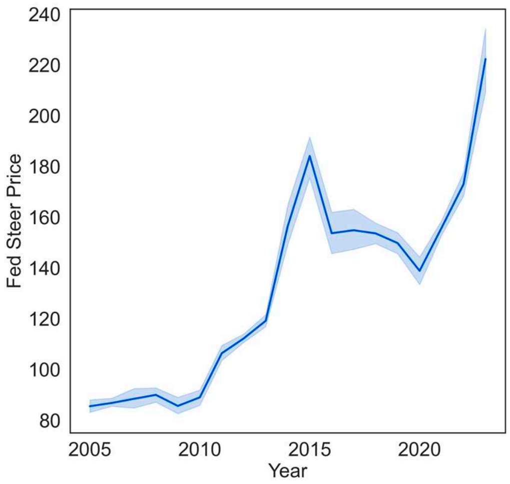


Figure 2. Alberta fed steer prices (Cdn$/cwt) from Jan. 2005 to Sep. 2023**.**

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 December 2023 doi:10.20944/preprints202312.1090.v1

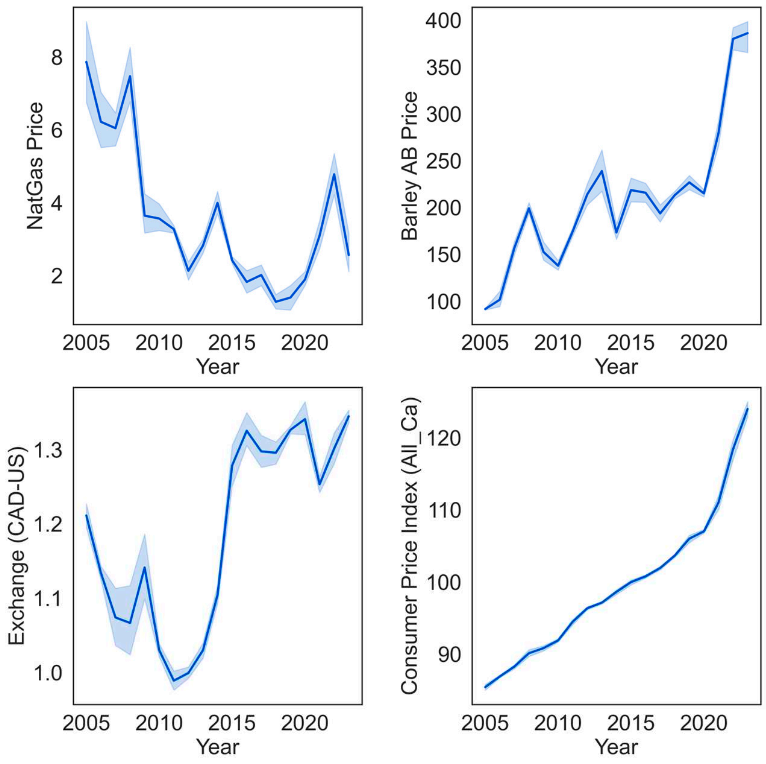
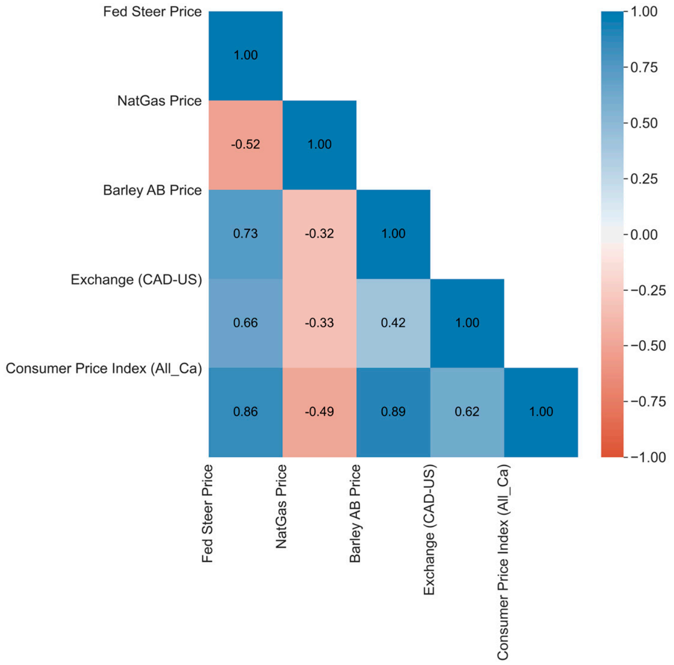


Figure 3. Time series variations of the predictors from Jan. 2005 to Sep. 2023**.**

We used Pearson correlation analysis to investigate the correlations between the predictors and fed steer prices. The correlation values are statistically significant (p-value < 0.01) between fed steer and predictor variables and are displayed in Figure 4. It shows that there is a high positive relationship between fed steer cattle price and the Canadian consumer price index (0.86), Alberta barley prices (0.73), and the exchange rate (0.66). Cattle prices, on the other hand, are negatively correlated with the natural gas price (-0.52).



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 December 2023 doi:10.20944/preprints202312.1090.v1

Figure 4. Correlation between fed steer price and predictors (p-values < 0.01).

3.3. Data Preprocessing & Partitioning (Train - Test) & Tunning

First, the quality of the data is visually checked by searching for obvious errors and outliers and missing data. Then we estimate the multivariate models using the training dataset on the scaled trained data (min-max scaler). Subsequently, we utilized these models to make predictions and assessed their performance on the scaled test dataset. We used 80% of the data for training and 20% for testing. The next step is to 'tune' the machine learning model, which is a way of prioritizing different features in their importance in the overall prediction performance. We used the 'hyper-parameter module technique' to tune our training datasets. This technique discovers the ideal hyper-parameter for each specific machine learning algorithm individually by comparing several model settings and comparing the metric to get the best combination of settings. The tune model hyper-parameter approach is used to improve the performance of the model [34].

In univariate modeling, we first assessed the time-series model's stability and performance over time using rolling-window analysis. The rolling average window is a technique used in analyzing time-series data. It implies calculating the average of the window size, which is a predetermined number of consecutive data points. Once the window size is determined, the model is estimated and the information criteria are assessed to determine the optimal smoothing window width. The machine learning feature engineering in data preprocessing selected a three-month window as optimal while also considering agricultural production seasonality, which is an important aspect of livestock marketing in Canada. Therefore, we run a rolling average over a three-month window for monthly timeseries on the fed steer prices. Also the fixed 80-20 split was deemed appropriate for the univariate ML forecasting modeling.

In univariate time series modeling, as in standard econometric approaches, we need to test the stationarity assumption before determining the machine learning parameters of the ARIMA, SARIMA, and SARIMAX models. We used Augmented Dickey Fuller (ADF) tests and determined that first order differentiation was sufficient to transform the Alberta fed steer price series to become stationary.

3.4. An Introduction to Machine Learning Algorithms and Description

Despite the limited applications in agricultural commodity price analysis, machine learning is widely used in a variety of other fields to address complex problems that are difficult to solve with traditional analytical methods [35]. ML is an artificial intelligence branch that uses algorithms rather than model-based analysis [36] to systematically synthesize the core connections between data and information with the purpose of predicting future scenarios. The main strength of machine learning is in identifying underlying relationships within datasets via pattern detection and prediction. ML systems can also detect disruptions to existing models and redesign and retrain themselves to adapt to and coevolve with new information. By relying on historical experience, the machine learning process plays a critical role in generalizing prediction problems to allow for maximum extraction of useful information from prior observed behaviors and patterns. Thus, historical observed data become 'training' datasets for the machine learning algorithms and better allow the ML model to generate largely accurate predictions even in novel situations. Many big data applications use ML to run at optimal efficiency. Here, we applied our ML techniques to analyze fed steer prices using two different approaches; multivariate and univariate modeling.

3.4.1. Multivariate Analysis

After preparing the data matrix, then we applied multivariate and univariate algorithms to predict Alberta fed steer prices. For multivariate machine learning regression modeling, we applied three robust and widely used algorithms; random forest, adaboost, and support vector machines.

Support vector machines (SVM) is a commonly used classification technique that properly categorizes data. Theoretically, it only takes a short training sample, and is unaffected by the number of dimensions but can be computationally intensive. Furthermore, effective approaches for training SVM are being developed at a swift pace and they can also be used for regression purposes by making minor changes [37–39].

Random forests, which were introduced by Breiman [40], are a set of tree predictors in which each tree is determined by the values of a random vector selected separately with an identical distribution of the trees in the forest. As a widely used classification and regression approach, random forest has proved quite an effective method, which aggregates numerous randomized decision trees and averages their predictions. It has proved able to perform well in situations where the number of variables exceeds the number of observations. Furthermore, it is adaptable to a variety of unstructured learning tasks and provides measures of variable significance, making it suitable for large-scale problems [41]. The RF algorithm assesses the significance of every feature in the prediction process, and displaying lower sensitivity to feature scaling and normalization. This characteristic makes it simpler for training and tuning.

The Adaboost algorithm or adaptive boosting is another multivariate method that we applied in this study. Adaboost, among the initial practical boosting techniques, was pioneered by Freund and Schapire [42]. Its primary concept is based on merging multiple classifiers, termed weak learners, into a singular classifier called strong classifier, by optimizing it through a weighted linear combination and integrating one weak classifier at each step.

Boosting is an ensemble method and employs multiple predictors to enhance accuracy in regression and classification tasks. To amplify and diversify the training dataset, boosting involves sequential sampling, repeatedly drawing samples with replacement from the original data. These methods are learned in a series, primarily benefiting from unstable learners like neural networks or decision trees. There's some indication that boosting leads to heightened accuracy levels [43,44].

3.4.2. Univariate Analysis

For univariate approach we used the autoregressive integrated moving average (ARIMA) model, seasonal ARIMA (SARIMA), and the seasonal autoregressive integrated moving average with exogenous factors (SARIMAX). Here, to predict fed steer prices, we only used its previous or historical time series data. Univariate time-series analysis is a method for explaining sequential problems over regular time intervals. When a continuous variable is time-dependent, it is advantageous to apply this method especially in finding consistent patterns in market data.

ARIMA is a class of models that explains a time series based on its own past values. ARIMA models can be used to model any non-seasonal time series that has patterns and isn't random white noise. Making the time series stationary is the first step in creating an ARIMA model, which is achieved through differencing. Depending on the complexity of the series, multiple levels of differencing may be required. Linear regression machine learning models work best when the predictors are not correlated and are independent of one another.

The problem with the basic ARIMA model is that it does not account for seasonality. Considering the seasonality effect, seasonal terms should be added to the ARIMA model to create the Seasonal ARIMA model (SARIMA). Seasonal differencing is used by SARIMA which is similar to regular differencing, except that instead of subtracting consecutive terms, the value from the previous season is subtracted.

The SARIMAX model is able to deal with external factors. We can include an external predictor, also known as an 'exogenous variable' in the model with the seasonal index. The seasonal index repeats every frequency cycle.

Table 3 presents a summary of the multivariate and univariate machine learning algorithms that were assessed in this study.

Table 3. Summary description of the multivariate and univariate ML approaches used in this study.

Modeling	Acronym	Algorithm	Description
Multivariate Approaches	RF	Random Forest Regression	RFR is an integrated learning method, a general-purpose and quite effective classification and regression

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 December 2023 doi:10.20944/preprints202312.1090.v1

			approach. It is a technique that ensembles numerous randomized decision trees and averages their predictions.
	AB	Adaboost Regressor	AdaBoost stands as a widely used classification algorithm. Throughout the training process, the sample's distribution weight is enhanced as the error rate rises; conversely, it diminishes, as the new distribution weight decreases. Subsequently, samples are continuously trained based on these altered distribution weights. The objective is to yield robust results by minimizing subsequent model errors, ultimately achieving higher accuracy rates [45,46].
	SVM	Support Vector Machines	SVM is a supervised-learning strategy that uses a symmetrical loss function that penalizes both high and low misestimates equally and it has been shown to be an effective method for estimating real-value functions [47]. It has the capability to conduct both linear and non-linear classification and regression. However, dealing with large datasets can be a challenging task [48].
Univariate Approaches	ARIMA	Auto Regressive Integrated Moving Average	ARIMA is a modeling algorithm based on the idea that past values of a time series can be used to predict future values by themselves, also taking into account autocorrelation in the error terms and stationarity.
	SARIMA	Seasonal Auto Regressive Integrated Moving Average	SARIMA is defined as the 'Seasonal' ARIMA model, and it is formed by adding seasonal lag and moving average terms to an ARIMA model.
	SARIMAX	Seasonal Auto Regressive Integrated Moving Average with exogenous factors	SARIMAX model is another form of SARIMA model with an external predictor, also known as an exogenous variable e.g. seasonal index.

3.5. Validation Methods

A comparison between the multivariate and univariate algorithms are done to evaluate the best models' performance on cattle price prediction for verification of the forecasts. To minimize errors in prediction models, predicted prices are assessed using mean absolute error (MAE), root mean square error (RMSE), and mean square error (MSE) [49–52].

The mean absolute error (MAE) is an extensively used metric for verifying a deterministic prediction and shows the magnitude of the error regardless of the prediction value [53,54]. The average distance between a data point and the fitted line, measured along a vertical line, is known as the root mean squared error (RMSE). RMSE is sensitive to outliers and exhibits both under- and over-estimation in the same pattern. The mean squared error (MSE) measures the average squared gap between observed and predicted values. By utilizing squared units rather than the original data units, it magnifies the influence of larger errors, causing them to be penalized more heavily than smaller errors. This attribute is crucial when identifying a model with smaller errors.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 December 2023 doi:10.20944/preprints202312.1090.v1

$$\mathbf{MAE} = \frac{\mathbf{1}}{\mathbf{n}}\sum_{\mathbf{i=1}}^{\mathbf{n}}|\mathbf{O_i} - \mathbf{P_i}|$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(O_i - P_i)^2}$$

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(O_i - P_i)^2$$

where, $\mathbf{O_i}$ is observed data and $\mathbf{P_i}$ is a deterministic prediction and $\mathbf{n}$ is the number of observations. Models with the lowest MAE, RMSE, and MSE were assumed to be the best models, as these metrices are negatively oriented.

4. Results and Discussions

4.1. Feature Selection & Hyperparameter Tuning

Fine-tuning hyperparameters is crucial for algorithms and enhancing the machine learning model's overall performance. This process is established before the learning phase and is conducted externally to the model. Without adequate hyperparameter tuning, the model can’t deliver accurate results since the minimization of the loss function does not take place and will be prone to make errors. Hyperparameter tuning aims to identify the optimal values that maximize the model's performance, minimize loss, and generate improved modeling results. Table 4 summarizes the most important parameters in multivariate ML models, with the best selected hyperparameters for our data shown. Tables 5 and 6 also show the best selected hyperparameters of univariate models and the best selected time series models residuals using the results.

Table 4. Summary statistics of multivariate ML models best selected hyperparameters.

RF hyperparameter	RF best params	SVR hyperparameter	SVR best params	AB hyperparameter	AB best params
Bootstrap	True	C	1	Base estimator max depth	10
Ccp alpha	0.0	Epsilon	0.1	Learning rate	1
Criterion	mse	Kernel	rbf	N estimators	300
Min impurity decrease	0.0	-	-	-	-
Min samples leaf	1	-	-	-	-
Min samples split	2	-	-	-	-
Min weight fraction leaf	0.0	-	-	-	-
N estimators	300	-	-	-	-
N jobs	1	-	-	-	-
Max features	auto	-	-	-	-
Max samples	None	-	-	-	-
Max leaf nodes	None	-	-	-	-
Max depth	None	-	-	-	-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 December 2023 doi:10.20944/preprints202312.1090.v1

Table 5. Summary statistics of univariate ML models best selected hyperparameters**.**

Hyperparameter	ARIMA	SARIMA	SARIMAX
start_p	1	1	1
start_q	1	1	1
max_p	3	3	3
max_q	3	3	3
m	1	12	12
test	adf	adf	adf
seasonal	False	False	True
trace	True	True	True
start_P	0	0	0
D	0	1	1
stepwise	True	True	True

In the univariate time series models, we applied a stepwise approach to find the best model with the lowest Akaike information criterion (AIC) among multiple combinations of model parameters for the ARIMA, SARIMA, and SARIMAX models. Table 6 describes some additional summary statistics of time series models residuals using the results. Prob (Q) is the p-value for the null hypothesis that the residuals have no correlation structure. Prob (JB) is the p-value associated with the null hypothesis that the residuals are Gaussian normally distributed. If either of the p-values is less than 0.05 the hypothesis is rejected. When the values of prob (Q) and prob (JB) are near 0.00, we can reject the null hypothesis that the residuals are not Gaussian normally distributed and have some correlation structure. The Akaike in-formation criterion (AIC) and Bayesian information criterion (BIC) were used to estimate the parameters of the best fit ARIMA model in time series modeling [55]. Using the stepwise approach, a seasonal interval of 12 months or one year is chosen for the SARIMA based machine learning problem, which implies that price patterns repeat in a similar way every year, rather than at a higher frequency pattern every quarter or season. Within SARIMAX, the seasonal index is chosen as an exogenous term due to its repetitive nature across each frequency cycle. In this case, we applied a seasonal decomposition employing the multiplicative model over a span of 36 months.

Table 6. Summary statistics of residuals for the best selected univariate ML models**.**

Hyperparameter	ARIMA	SARIMA	SARIMAX
Best Model	ARIMA(3,1,1) (0,0,0)[0]	ARIMA(0,1,3)(0,1,1)[12]	ARIMA(2,0,0)(1,1,2)[12]
AIC	960.02	878.99	1000.73
BIC	980.46	895.75	1027.63
Prob (Q)	0	0.09	0
Prob (JB)	0	0	0.56
Heteroske-dasticity (H)	2.51	3.07	2.39
Ljung-Box (Q)	70.43	52.62	183.15
Skew	-0.62	-0.31	-0.02

4.2. Validation of Multivariate and Univariate Machine Learning Models

Upon running the ML models on the training data, we computed the mean absolute error (MAE), root mean square error (RMSE), and mean squared error (MSE) to assess the models performance using the remaining 20% of the test dataset. Smaller values of MAE, RMSE, and MSE indicate better performance since they are negatively oriented metrics. As a result, we can be confident in the models ability to accurately predict future cattle price values. The accuracy of our forecast models are summarized in Table 7 and illustrated in Figure 5.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 December 2023 doi:10.20944/preprints202312.1090.v1

Comparing the performance of all predicting models, the random forest and adaboost models, were the selected as outstanding models according to the validation metrices on the test data for predicting monthly Alberta fed steer cattle prices.

Table 7. The accuracy metrics of the ML price predicting models on the test data.

Model	ML Algorithms	MAE	RMSE	MSE
Multivariate	RF	0.19	0.28	0.08
	AB	0.19	0.28	0.08
	SVR	0.33	0.43	0.19
Univariate	ARIMA	3.24	4.35	18.89
	SARIMA	1.76	2.18	4.76
	SARIMAX	2.50	3.05	9.31

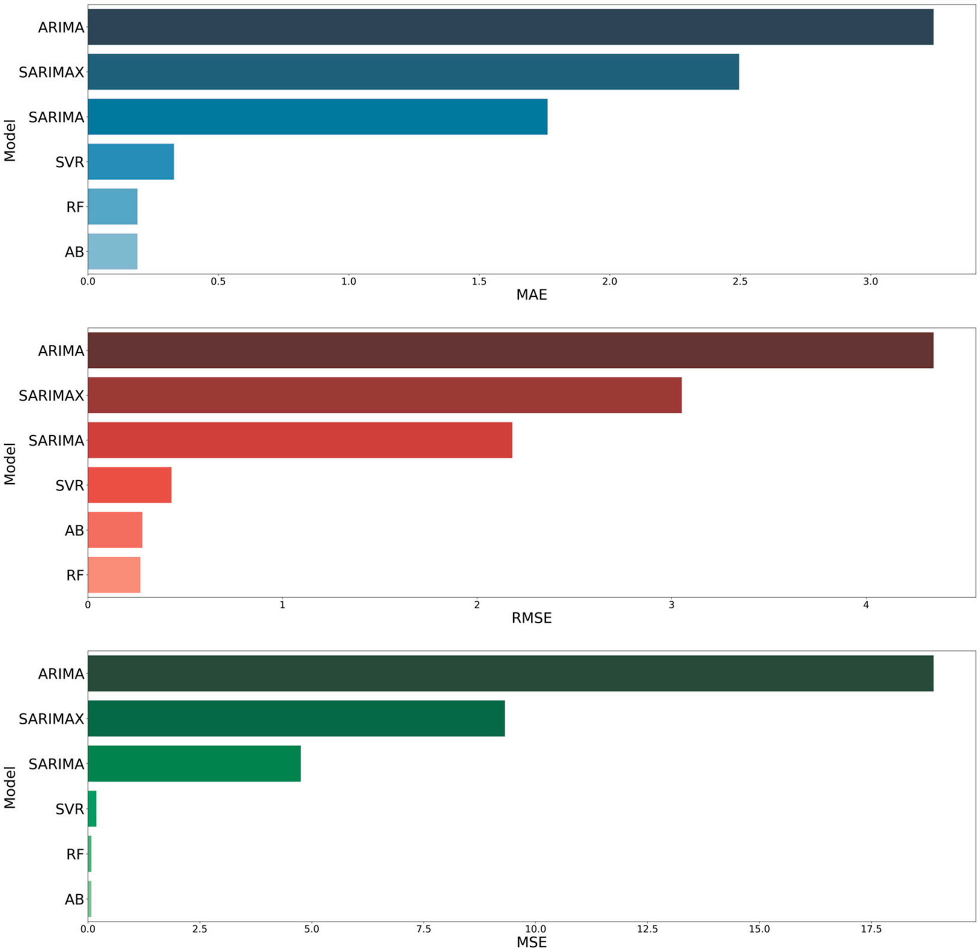


Figure 5. Visualization of ML models performances for predicting fed steer cattle prices.

4.3. Applying Probabilistic Modeling Approach to the Selected Machine Learning Model

Since both RF and AB demonstrated comparable high performance, we opted to present the remaining validation visualizations solely using the RF model. The RF model's outcome estimates the anticipated cattle price but lacks variance, leading to a decrease in variance for the RF deterministic forecasts. The most statistically valid method to produce predictions with realistic variance is modeling the probability density of the residual and taking it into account. The premise is that the scaled actual price and residuals are realizations of a univariate Gaussian random variable,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 December 2023 doi:10.20944/preprints202312.1090.v1

with an average of zero and a variance based on the disparities between the actual and RF-predicted values. This approach yields realizations of the conditional distribution of the actual price values using a probabilistic approach. Figure 6 shows the visualization of the scaled actual and probabilistic RF predicted for the test data set from January 2005 to September 2023. Also, the boxplots illustrate 500 realizations of the probabilistic generated prices indicating the inner quartile range within these forecasts. The whiskers represent the 0.025 and 0.975 quantiles. The graph demonstrates that the outcomes of our probabilistic RF prediction model fall within the 95% confidence interval. The zoomed visualization of the test data verification (the right side of the Figure 6) is represented in Figure 7.

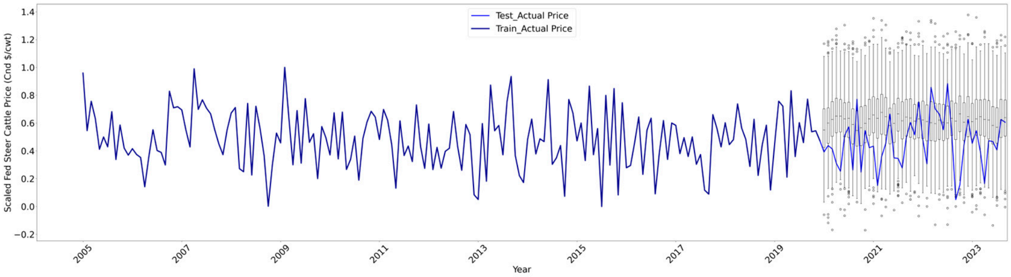


Figure 6. Actual and probabilistic RF predicted cattle price by RF model on test data set from 2005 to 2023. Boxplot shows 500 realizations of the probabilistic RF model. The boxes display the median and interquartile range, while the whiskers indicate the 95% interval.

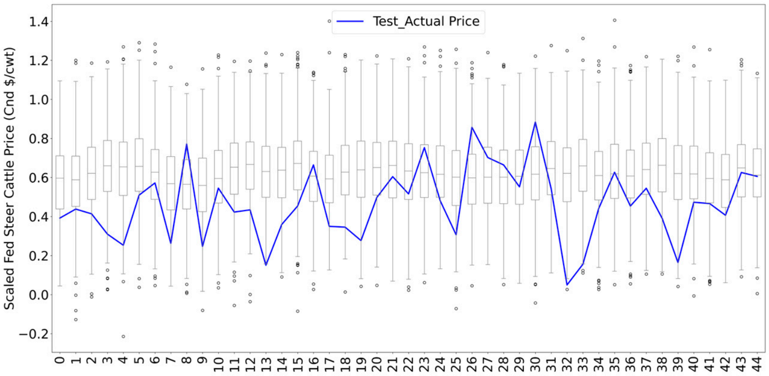


Figure 7. Zoomed graph of actual and probabilistic RF predicted cattle price by RF model on test data set**.**

5. Conclusions

Alberta's cattle industry is important to the province's economy. The province is a key producer of beef in Canada, boasting the largest distribution of feeder beef cattle and calves. Cattle prices in Alberta have experienced considerable fluctuations from 2005 till present due to multiple influencing factors. We believe that forecasting these prices holds significance for farmers, policymakers, insurance companies and traders, enabling informed marketing decisions and risk management. This study highlights the successful predictive ability of machine learning algorithms in anticipating cattle prices and their associated uncertainties in Alberta. Factors like data quality, variable selection, and model optimization are pivotal aspects without compromising performance. Greater use of standard machine learning methods in commodity price modeling has the potential to improve our ability to analyze agricultural commodity prices. In this paper, we have conducted a side-by-side comparison of univariate and multivariate machine learning modeling methods on monthly fed steer prices in Alberta between January 2005 and September 2023. We can summarize this study's outcomes with three main points:

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 December 2023 doi:10.20944/preprints202312.1090.v1

1. Multivariate modeling, incorporating additional key variables as predictors, demonstrated a notable advantage over univariate approaches in our investigation.
2. Probabilistic modeling has an advantage compared to deterministic modeling. By employing probabilistic modeling, we consider uncertainties and incorporate them into deterministic RF predictions, providing a more realistic context for the predicted values with the selected RF model. This process should be more routinely applied to other machine learning modeling studies.
3. Lastly, in the comparison between ML algorithms, Adaboost and Random Forest models showed similar and robust validation performance concerning our variables.

Machine learning methods with additional probabilistic modeling can thus improve on our analytical capacity by enabling more accurate model selection by 'letting the data speak' as well as better incorporating new information and complex time series data structure elements. Future steps involve investigating whether a multivariate approach could further enhance the prediction of fed steer prices in a year characterized by significant shocks and dynamic influences such as inflation, more variation on grain prices and exchange rates.

Author Contributions: E.R. and E.S. contributed equally to the writing of the original draft of this manuscript. Conceptualization, E.R., M.K. and E.S..; methodology, E.R., M.K. and E.S..; software, E.R. and M.K. ; validation, E.R., M.K. and E.S.; formal analysis, E.R., and E.S.; writing—original draft preparation, E.R. and E.S.; writing—review and editing, E.R, M.K. and E.S.; visualization, E.R.; supervision, E.S. All authors have read and agreed to the published version of the manuscript.

Funding: No external funding was used to support this research.

Data Availability Statement: The monthly Alberta fed steer prices used in this research are proprietary data for CanFax Research Service subscribers but may be available on request. The monthly average Alberta fed steer prices are freely available via Statistics Canada (Table 32-10-0077-01 Farm product prices, crops and livestock, https://doi.org/10.25318/3210007701-eng).

Acknowledgments: Special thanks go to Brenna Grant, executive director of CanFax and CanFax Research Services in Calgary, Alberta, Canada for making the long-term Alberta fed steer weekly price data available for research.

Conflicts of Interest: The authors declare no conflict of interest and CanFax had no role in the design of the study; in the analyses, or interpretation of data; in the writing of the manuscript. Permission was requested and received from CanFax to publish our research findings.

References

1. Liakos, K.G.; Busato, P.; Moshou, D.; Pearson, S.; Bochtis, D.J.S. Machine learning in agriculture: A review. **2018**, *18*, 2674.
2. Ouyang, H.; Wei, X.; Wu, Q.J.J.o.A.E. Agricultural commodity futures prices prediction via long-and short-term time series network. **2019**, *22*, 468-483.
3. Chen, Z.; Li, C.; Sun, W. Bitcoin price prediction using machine learning: An approach to sample dimension engineering. *Journal of Computational and Applied Mathematics* **2020**, *365*, doi:10.1016/j.cam.2019.112395.
4. Rahmani, E.; Friederichs, P.; Keller, J.; Hense, A. Development of an effective and potentially scalable weather generator for temperature and growing degree days. *Theoretical and Applied Climatology* **2016**, *124*, 1167-1186.
5. Colino, E.V.; Irwin, S.H.J.A.J.o.A.E. Outlook vs. futures: Three decades of evidence in hog and cattle markets. **2010**, *92*, 1-15.
6. Marsh, J.M. US feeder cattle prices: effects of finance and risk, cow-calf and feedlot technologies, and Mexican feeder imports. *Journal of Agricultural Resource Economics* **2001**, 463-477.
7. Tomek, W.G. Commodity prices revisited. *Agricultural Resource Economics Review* **2000**, *29*, 125-137.
8. Zhao, H. Futures price prediction of agricultural products based on machine learning. *Neural Computing Applications* **2021**, *33*, 837-850.
9. Blank, S.C.; Saitone, T.L.; Sexton, R.J.J.J.o.A.; Economics, R. Calf and yearling prices in the western United States: Spatial, quality, and temporal factors in satellite video auctions. **2016**, 458-480.
10. Sanders, D.R.; Manfredo, M.R.J.J.o.A.; Economics, R. USDA livestock price forecasts: A comprehensive evaluation. **2003**, 316-334.
11. Oliveira, R.A.; O'connor, C.W.; Smith, G.W. Short-run forecasting models of beef prices. *Western Journal of Agricultural Economics* **1979**, 45-55.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 December 2023 doi:10.20944/preprints202312.1090.v1

12. Zapata, H.O.; Garcia, P.J.W.J.o.A.E. Price forecasting with time-series methods and nonstationary data: An application to monthly US cattle prices. **1990**, 123-132.
13. Linnell, P.B. Forecasting Fed Cattle Prices: Errors and Performance During Periods of High Volatility. Colorado State University, 2017.
14. Kohzadi, N.; Boyd, M.S.; Kermanshahi, B.; Kaastra, I.J.N. A comparison of artificial neural network and time series models for forecasting commodity prices. **1996**, *10*, 169-181.
15. Jeong, S.; Ko, J.; Yeom, J.M. Predicting rice yield at pixel scale through synthetic use of crop and deep learning models with satellite data in South and North Korea. *Sci Total Environ* **2022**, *802*, 149726, doi:10.1016/j.scitotenv.2021.149726.
16. Sharma, A.; Jain, A.; Gupta, P.; Chowdary, V. Machine Learning Applications for Precision Agriculture: A Comprehensive Review. *IEEE Access* **2021**, *9*, 4843-4873, doi:10.1109/access.2020.3048415.
17. Liu, W.; Shao, X.-F.; Wu, C.-H.; Qiao, P. A systematic literature review on applications of information and communication technologies and blockchain technologies for precision agriculture development. *Journal of Cleaner Production* **2021**, *298*, doi:10.1016/j.jclepro.2021.126763.
18. Maroli, A.; Narwane, V.S.; Gardas, B.B. Applications of IoT for achieving sustainability in agricultural sector: A comprehensive review. *J Environ Manage* **2021**, *298*, 113488, doi:10.1016/j.jenvman.2021.113488.
19. Meshram, V.; Patil, K.; Meshram, V.; Hanchate, D.; Ramkteke, S.D. Machine learning in agriculture domain: A state-of-art survey. *Artificial Intelligence in the Life Sciences* **2021**, *1*, doi:10.1016/j.ailsci.2021.100010.
20. Chen, Z.; Goh, H.S.; Sin, K.L.; Lim, K.; Chung, N.K.H.; Liew, X.Y. Automated Agriculture Commodity Price Prediction System with Machine Learning Techniques. *Advances in Science, Technology and Engineering Systems Journal* **2021**, *6*, 376-384, doi:10.25046/aj060442.
21. Tian, H.; Wang, P.; Tansey, K.; Zhang, J.; Zhang, S.; Li, H. An LSTM neural network for improving wheat yield estimates by integrating remote sensing data and meteorological data in the Guanzhong Plain, PR China. *Agricultural and Forest Meteorology* **2021**, *310*, doi:10.1016/j.agrformet.2021.108629.
22. Divisekara, R.W.; Jayasinghe, G.; Kumari, K.J.S.B.; Economics. Forecasting the red lentils commodity market price using SARIMA models. **2021**, *1*, 1-13.
23. Sharma, R.; Kamble, S.S.; Gunasekaran, A.; Kumar, V.; Kumar, A.J.C.; Research, O. A systematic literature review on machine learning applications for sustainable agriculture supply chain performance. **2020**, *119*, 104926.
24. Kamir, E.; Waldner, F.; Hochman, Z. Estimating wheat yields in Australia using climate records, satellite image time series and machine learning methods. *ISPRS Journal of Photogrammetry and Remote Sensing* **2020**, *160*, 124-135, doi:10.1016/j.isprsjprs.2019.11.008.
25. Yamaç, S.S.; Todorovic, M. Estimation of daily potato crop evapotranspiration using three different machine learning algorithms and four scenarios of available meteorological data. *Agricultural Water Management* **2020**, *228*, doi:10.1016/j.agwat.2019.105875.
26. van Klompenburg, T.; Kassahun, A.; Catal, C. Crop yield prediction using machine learning: A systematic literature review. *Computers and Electronics in Agriculture* **2020**, *177*, doi:10.1016/j.compag.2020.105709.
27. Vidyarthi, S.K.; Tiwari, R.; Singh, S.K.; Xiao, H.W. Prediction of size and mass of pistachio kernels using random Forest machine learning. *Journal of Food Process Engineering* **2020**, *43*, doi:10.1111/jfpe.13473.
28. Cai, Y.; Guan, K.; Lobell, D.; Potgieter, A.B.; Wang, S.; Peng, J.; Xu, T.; Asseng, S.; Zhang, Y.; You, L.; et al. Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. *Agricultural and Forest Meteorology* **2019**, *274*, 144-159, doi:10.1016/j.agrformet.2019.03.010.
29. Kouadio, L.; Deo, R.C.; Byrareddy, V.; Adamowski, J.F.; Mushtaq, S.; Phuong Nguyen, V. Artificial intelligence approach for the prediction of Robusta coffee yield using soil fertility properties. *Computers and Electronics in Agriculture* **2018**, *155*, 324-338, doi:10.1016/j.compag.2018.10.014.
30. Prajapati, B.P.; Kathiriya, D.R.J.I.J.o.C.A. Towards the new Similarity Measures in Application of Machine Learning Techniques on Agriculture Dataset. **2016**, *156*.
31. Shepherd, A. Market information services: Theory and practice; Food & Agriculture Org.: 1997.
32. Dorward, A. Agricultural labour productivity, food prices and sustainable development impacts and indicators. *Food policy* **2013**, *39*, 40-50.
33. CanadaBeef. Canada's Beef Industry FAST FACTS; CANADA BEEF, 2021.
34. Jumin, E.; Basaruddin, F.B.; Yusoff, Y.B.M.; Latif, S.D.; Ahmed, A.N. Solar radiation prediction using boosted decision tree regression model: A case study in Malaysia. *Environ Sci Pollut Res Int* **2021**, *28*, 26571-26583, doi:10.1007/s11356-021-12435-6.
35. Maulud, D.; Abdulazeez, A.M. A Review on Linear Regression Comprehensive in Machine Learning. *Journal of Applied Science and Technology Trends* **2020**, *1*, 140-147, doi:10.38094/jastt1457.
36. Athey, S.; Imbens, G.W.J.A.R.o.E. Machine learning methods that economists should know about. **2019**, *11*, 685-725.
37. Cortes, C.; Vapnik, V.J.M.l. Support-vector networks. **1995**, *20*, 273-297.
38. Sharifzadeh, M.; Sikinioti-Lock, A.; Shah, N. Machine-learning methods for integrated renewable power generation: A comparative study of artificial neural networks, support vector regression, and Gaussian

Process Regression. *Renewable and Sustainable Energy Reviews* **2019**, *108*, 513-538, doi:10.1016/j.rser.2019.03.040.

39. Wu, X.; Kumar, V.; Ross Quinlan, J.; Ghosh, J.; Yang, Q.; Motoda, H.; McLachlan, G.J.; Ng, A.; Liu, B.; Yu, P.S.; et al. Top 10 algorithms in data mining. *Knowledge and Information Systems* **2007**, *14*, 1-37, doi:10.1007/s10115-007-0114-2.
40. Breiman, L. Random forests. **2001**, *45*, 5-32.
41. Biau, G.; Scornet, E. A random forest guided tour. *Test* **2016**, *25*, 197-227, doi:10.1007/s11749-016-0481-7.
42. Freund, Y.; Schapire, R.; Abe, N. A short introduction to boosting. *Journal-Japanese Society For Artificial Intelligence* **1999**, *14*, 1612.
43. Drucker, H. Improving regressors using boosting techniques. In Proceedings of the Icml, 1997; pp. 107-115.
44. Kummer, N.; Najjaran, H. Adaboost. MRT: Boosting regression for multivariate estimation. *Artif. Intell. Res.* **2014**, *3*, 64-76.
45. Lu, J.; Hu, H.; Bai, Y. Generalized radial basis function neural network based on an improved dynamic particle swarm optimization and AdaBoost algorithm. *Neurocomputing* **2015**, *152*, 305-315.
46. Sevinç, E. An empowered AdaBoost algorithm implementation: A COVID-19 dataset study. *Computers & Industrial Engineering* **2022**, *165*, 107912.
47. Awad, M.; Khanna, R. Efficient learning machines: theories, concepts, and applications for engineers and system designers; Springer nature: 2015.
48. Greener, J.G.; Kandathil, S.M.; Moffat, L.; Jones, D.T. A guide to machine learning for biologists. *Nature Reviews Molecular Cell Biology* **2022**, *23*, 40-55.
49. Kumari, S.N.; Tan, A.J.T.S. Modeling and forecasting volatility series: with reference to Gold price. **2018**, *16*, 77-63.
50. Mahmoudzadeh, H.; Matinfar, H.R.; Taghizadeh-Mehrjardi, R.; Kerry, R. Spatial prediction of soil organic carbon using machine learning techniques in western Iran. *Geoderma Regional* **2020**, *21*, doi:10.1016/j.geodrs.2020.e00260.
51. Rahmani, E. The effect of climate variability on wheat in Iran. 2015.
52. Vijh, M.; Chandola, D.; Tikkiwal, V.A.; Kumar, A.J.P.C.S. Stock closing price prediction using machine learning techniques. **2020**, *167*, 599-606.
53. Dubey, A.K.; Kumar, A.; García-Díaz, V.; Sharma, A.K.; Kanhaiya, K.J.S.E.T.; Assessments. Study and analysis of SARIMA and LSTM in forecasting time series data. **2021**, *47*, 101474.
54. Gneiting, T.; Raftery, A.E.; Westveld III, A.H.; Goldman, T. Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Monthly Weather Review* **2005**, *133*, 1098-1118.
55. Nokeri, T.C. Forecasting Using ARIMA, SARIMA, and the Additive Model. In *Implementing Machine Learning for Finance*; Springer: 2021; pp. 21-50.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.