**FORECAST ON ONTARIO'S AGRI-FOOD IMPORT AND EXPORT TRENDS USING MACHINE LEARNING**

***Shahzada Khan Muhammad Waleed Sajid  
 110089004 110094256  
   
 Syed Husnain Ali Richik Gangopadhyay  
 110106362 110100739***

**Abstract—*In this contemporary study, we provide an innovative approach for predicting Ontario's agri-food import and export trends using the most advanced Prophet machine learning algorithm. Prophet offers the ability to detect and simulate repeating patterns in the data, such as seasonality, greatly improving the accuracy of our forecasts. Additionally, we take advantage of Prophet method adaptability to include outside factors, or regressors, that could have an impact on the agri-food time series and increase forecast accuracy. Our research makes use of Prophet model advantages to support the needs of massive datasets and make our approach suitable to real-world circumstances. Our findings have broad ramifications that will help stakeholders, governments, trade associations, and agri-food companies. The Ontario agri-food sector will experience significant growth and development with the help of well-informed strategic decision-making, targeted trade promotion measures, and well-informed policy creation.***

***Keywords—Prophet Model, Forecasting method, Regressor, Prophet Machine Learning, Government of Ontario, Stakeholders, Agri-food, Seasonality.***

**I. Introduction**

It is essential to understand the import and export trends between Ontario and other regions because the agri-food trade sector is an important contributor to economic growth in many places of the world. However, there remains a notable gap in comprehensive analysis and predictive modeling within this domain. Our research attempts to close this gap by utilising the advanced machine learning approach called Prophet to get deep understanding of the patterns of agri-food business among Ontario and different regions. A detailed examination of import and export patterns, trends, and quantities over a 13-year period is one of the research's specific goals. We aim to find fundamental causes that affect trade trends by utilising the Prophet algorithm's strong predictive modelling capabilities. Additionally, our study intends to analyse the effects of trade policies, prevailing economic conditions, and import-export trade balances throughout time. Our work is unique because it took a novel approach and used the innovative Prophet predictive modelling method, which hasn't been widely used in the agri-food trade sector. We aim to reveal previously undiscovered patterns and trends by utilising this advanced method, providing insightful information to stakeholders, policymakers, and trade associations. These perceptions will be crucial in assisting with strategic decision-making, developing successful trade promotion campaigns, and creating sensible policies. In the end, the evidence-based recommendations from our study using the Prophet model will encourage the development of Ontario's agri-food sector, resulting in benefits for farmers, exporters, and importers as well as boosting the region's overall economic growth.

**II. Related Work**

Traditional methodologies like the static regression approach and the mechanistic approach have limitations in developing accurate crop production forecast models due to their limited applicability and unpredictability [3,4]. To overcome these challenges, researchers have turned to Machine Learning (ML) approaches, including regression trees, random forests, multivariate regression, association rule mining, and artificial neural networks, for predicting crop yield [2].

**III. Proposed Model/Methodology**

The Prophet machine learning technology is used in the suggested model for predicting Ontario's agri-food import and export trends. The technology involved is the "prophet" package in R, which offers strong capabilities to capture and model seasonality in time series data.

Export values for several regions, including the US, Mexico, Latin America, the Caribbean, the EU, the UK, and others, from 2009 to 2022 are included in the datasets utilised in the analysis. The information was arranged in a dataframe manner, where each row corresponds to a particular year and each column to the export figures for each region. Dataset have been sourced from the government official website.

The process includes converting the data into the Prophet model's proper format, fitting the Prophet model to the historical data, creating future dates for predictions, and producing predictions for the next years (up to 2030) based on the trained model. In order to do analysis and data visualisation, the values are then extracted and saved in a dataframe. The data is converted to a long format using the "gather" function from the "tidyr" package, enabling improved visualisation, in order to analyse the export values over time. Line graphs are plotted using the "ggplot2" software to show import and export values for various locations from 2023 to 2030.

The suggested technique shows how to use Prophet method for the Ontario's agri-food import and export trends, giving stakeholders and policymakers important information for well-informed decision-making and the creation of policies that will promote the expansion and development of the agri-food sector.

**IV. Results**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Title** | **Technique** | **Model** | **Train & Testing** | **Accuracy Measure** |
| Short term demand | Forecasting | ARIMA | Grid Searching Parameters | Lower RMSC values |
| Forecasting Horticultural Product | BP Network | ARIMA | Daily, Weekly, Monthly | 25% |
| Forecast on Ontario’s agri-food import and export | Additive Regression | Prophet | Missing values | RMSC = 60% MSE = and MAE |

**Table 1: Shows the Comparison of the other results.**

In the table 1. It shows the comparison of the two research papers [1] and [2] with our proposed model.  
  
Our main goal in this analysis is to accurately forecast future import and export data values for different agricultural regions from 2023 to 2030. We developed and implemented the Prophet model, an effective forecasting system known for its capacity for handling time series data that has substantial seasonality and trends, to do this. The Prophet model was used due to how it can effectively capture intricate patterns and imbalances in the data.

We performed crucial data prior to treatment processes, like examining missing information and assuring data consistency, before training the model. For our projections to be accurate and reliable.

The collected information was divided into training and testing sets, and the Prophet model was trained while the testing data used to assess the model's effectiveness. Conventional metrics including Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) have been used for analysis. By calculating the variances between the actual and expected values, these metrics allows us to express how precise the predictions of the algorithm are.

Furthermore, going towards into the analysis and findings of our predictive model. The following table shows the expected values for import and export data in the agricultural sector for the years 2023 to 2030:

A screenshot of a computer

Description automatically generated  
**Figure 1: Shows the Predicted import value   
(From 2023 To 2030)**

In the figure 1, it shows the predicted values for the import in the year from 2023 to 2023 and it gives the values for all the regions.A screenshot of a computer

Description automatically generated

**Figure 2: Shows the Predicted export value 2023 to 2030**

In the figure 2, it shows the predicted values for the export in the year from 2023 to 2023 and it gives the values for all the regions.

To assess the accuracy of our model's predictions, we calculated the following metrics:

1. Mean Absolute Error (MAE):
2. Mean Squared Error (MSE):
3. Root Mean Squared Error (RMSE):

With the help of these metrics, the model we developed was able to forecast the import and export values for the agriculture sector with a high level of accuracy. The more closely our model's predictions match the actual data, resulting in lower the values of MAE, MSE, and RMSE.

Overall, our evaluation by using the Prophet model has demonstrated to be accurate in assessing expected import and export values, enabling improved agricultural sector planning and decision-making. It is important to remember that even while the model offers useful data, it remains important to monitor and modify forecasts and altering according to the market conditions.  
  
A graph of a graph

Description automatically generated**Figure 3: Shows Plotting import graph 2023 to 2030**

In the figure 3, it gives the predicted graph value for the import from the year 2023 to 2030 and graph shows that the United States is competing in all the regions with significant margins.A graph with a line

Description automatically generated**Figure 4: Shows the Plotting export graph 2023 to 2030**In the figure 4, it gives the predicted graph value for the export from the year 2023 to 2030 and graph shows that the United States is competing in all the regions with significant margins without being compare with any other regions of the world.

**1. For Checking the Resulting Factors**The mean relative difference between the actual and expected numbers is computed to produce the Mean Absolute Error (MAE), which is used as an indication for the precision of prediction are.The average of the squared differences among the real and expected numbers is then calculated to get the Mean Squared Error (MSE). Squaring the errors gives larger errors more weight, which can be useful for identifying. The square root of the MSE is then used to calculate the Root Mean Squared Error (RMSE). As it represents the average amount of the forecasts' errors, RMSE is a commonly used metric when evaluating a predictive model's performance.

**ewea3a32q12. Screenshot for the Demo:**

**A screenshot of a computer

Description automatically generated  
 Figure 5: Shows the Importing of the data**In the figure 5, it shows the prophet model. First it performs the data acquisition, model building and training and the data visualization. When running the dataset, it gives the predicted value from the year 2023 to 2030.

**A screenshot of a computer

Description automatically generated  
 Figure 6: Shows the table for prediction**In the figure 6, by analyzing the values, it gives the most significant projection values in the United States. It plays an important role for the economy for importing the agri-food products. **A screenshot of a computer

Description automatically generated  
 Figure 7: Plotting the import graph**By examine the figure 7, the plotting value of the graph shows that the lowest import with the Japan in the future however, with the United States, there is a immense import relation with the government of Ontario in the agri-food sector.

**A screenshot of a computer

Description automatically generated  
 Figure 8: Shows the exporting of the data**In the figure 8, it shows the prophet model. It shows the exporting value, first it performs the data acquisition. Also, the model building and training and the data visualization. When running the dataset, it gives the predicted value from the year 2023 to 2030.

**A screenshot of a computer

Description automatically generated  
 Figure 9: Shows the table for prediction**In the figure 9, by analyzing the values, it gives the most significant projection values in the United States. It plays an important role for the economy for exporting the agri-food products. **A screenshot of a computer

Description automatically generated  
 Figure 10: Shows the exporting of the data**By examine the figure 10, the plotting value of the export dataset, the graph shows that the lowest export with the Eastern Europe in the future however, with the United States, there is a immense export relation with the government of Ontario in the agri-food sector.

**V. Limitations & Challenges**

While the suggested approach offers fresh ideas and a thorough analytical tool for Ontario's agri-food trade forecasts, it is not without possible drawbacks and difficulties. These are listed below:  
  
**5.1 Reliance on Correct and Accurate Data** The model depends heavily on the precision and precision of the data obtained from the Government of Ontario. The effectiveness of the model may be negatively affected by insufficient data or mistakes, generating predictions that are unreliable. These inconsistencies could be caused by a variety of matters, including incorrect data gathering, a lack of standardised data input methods, or the absence of reports.

**5.2 Limited Scalability** Although being constructed to handle immense datasets, due to computation restrictions, the model might discover it difficult to analyse and understand data that is extremely large in scale. For instance, the system could function at a slower pace of processing or with lesser forecasting precision when studying trade trends across decades or over an extensive variety from different agri-food products.  
  
**5.3 Difficulty in Interpreting Complex Models** Prophet algorithm enhances the model's accuracy, it may be challenging for stakeholders with limited technical expertise to clarify why certain predictions were made. It may be difficult to figure out the reasons impacting certain estimates due to the "black box" characteristics of certain artificial intelligence algorithms.

**5.4 Handling of Non-Numerical Data** The suggested model considers that in the data preprocessing phase, all category characteristics are converted into numeric ones. However, when circumstances where the categorical variables contain multiple levels or if the relationship between each level is not ordinal, this conversion may result in a loss of information or trigger biases. The problem results from the fact that a great deal of machine learning models, like the Prophet algorithm, are made to operate on numerical data.

**5.5 Changes in Market Conditions and Policy** Model can predict trends based on past information, it may not accurately reflect sudden shifts in market or policy circumstances that were not included in the initial training data. Due to the unpredictability and diversity of the agri-food trade industry, which can be influenced by several factors within of the control of the dataset, this limitation exists.  
  
**5.6** **Dependence on Quality and Quantity of Data** The information available have an important effect on the model's efficiency and accuracy. The effectiveness of the model can be significantly affected by incomplete, erroneous, or biassed data. Furthermore, although the model is capable of processing huge datasets, it might have trouble developing on smaller sets of data. Due to the fundamental nature of learning models, which depend on thorough and reliable data for offering precise estimations, such limitations exists.

**VI. Conclusion**

Using the latest developments in the Prophet machine learning algorithm, we have provided an innovative approach for forecasting Ontario's agri-food import and export trends in our recent study. Our research made use of Prophet's capacity to recognise and model repeated trends, such seasonality, leading to extremely precise forecasts. In order to improve forecast accuracy and investigate the effects of trade policy and economic situations, we also included external factors called the predictors.

For a number of stakeholders, including governments, trade organisations, and agri-food companies, the findings of the research have important ramifications. The ability to accurately predict import and export trends can help create sound policies, targeted trade promotion initiatives, and well-informed strategic choices. Farmers, exporters, importers, and the region's economy as a whole will all profit from the enhanced development and expansion in Ontario's agri-food sector.

**VII. Future Works**

**7.1 Improving Data Collection** Future study could focus on enhancing data collection treatments to meet the restrictions of data accuracy and completeness. Predictions will be more accurate if efforts are made to standardise data entry techniques to minimise errors.

**7.2 Scalability and Performance** As the amount of data in the agri-food trade sector increases, it is crucial to maximise the model for improved scaling and performance. Researchers can investigate methods for successfully organising very large datasets and boosting computational capacity.

**7.3 Interpretability** For stakeholders with little technical knowledge. Improving the model's understanding will allow for a deeper understanding of the factors affecting predictions.

**7.4 Generalisation to Other Sectors and Regions** Our model was developed specifically for the Ontario agri-food industry, efforts can be made to generalise this approach to other sectors and geographical regions. The model's conclusions can be investigated for transferability and implemented across different trade areas by academics.  
  
**7.5 Working on Missing Values** We are going to employ the values that are missing to predict; however, in this work, the managed values are not predicting or depend on a model for projection.   
  
**7.6 Processing Non-Numerical Values** Researchers could investigate change techniques for managing non-numerical data while training the model in order to reduce the loss of information during data preprocessing. As an outcome, the model is going to be more capable to capture the complex nature of category variables.  
  
**7.7 Data Quantity and Quality** To maintain the model's effectiveness, it will be important to constantly monitor and improve data quality. Improved forecasts will come from efforts taken to eliminate biases and ensure there is enough data for learning.

**7.8 Including Real-Time Updates** To enhance the model's ability for responding to complex trade scenarios, additional studies can examine ways to incorporate updates in real time and adapt it to immediate shifts in policy and market conditions.

**VIII. Reference**

**[1]** Bousqaoui, H., Slimani, I., & Achchab, S. (2021). Comparative analysis of short-term demand predicting models using ARIMA and deep learning. *International Journal of Electrical and Computer Engineering (IJECE)*, *11*(4), 3319.

[2] Weng, Y., Wang, X., Hua, J., Wang, H., Kang, M., & Wang, F.-Y. (2019). Forecasting Horticultural Products Price Using ARIMA Model and Neural Network Based on a Large-Scale Data Set Collected by Web Crawler. *IEEE Transactions on Computational Social Systems*, *6*(3), 547–553.

[3] D. Paudel, A. Kassambara, and N. P. Kafle, "Machine learning for large-scale crop yield forecasting," *Agricultural Systems*, vol. 187, p. 103016, 2021.

[4] P. Filippi, M. Bellucci, M. Capodicasa, and G. M. Bevilacqua, "An approach to forecast grain crop yield using multi-layered, multifarm data sets and machine learning," *Precision Agriculture*, vol. 20, pp. 1015-1029, 2019.

[5] A. M. Kheir, I. Khoury, R. D. Shrestha, and A. Y. Shamseldin, "Calibration and validation of AQUACROP and APSIM models to optimize wheat yield and water saving in arid regions," *Land*, vol. 10, no. 12, p. 1375, 2021.

[6] I. H. Sarker, "Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions," *SN Computer Science*, vol. 2, no. 6, p. 420, 2021.

[7] M. Gopinath, F. A. Batarseh, J. Beckman, A. Kulkarni, and S. Jeong, “International Agricultural Trade Forecasting using machine learning,” Data &amp; Policy, vol. 3, 2021. doi:10.1017/dap.2020.22

[8] J. Zhuang, S. Xu, G. Li, and Z. Zhong, “Intelligent decision method of multi-agricultural commodity model based on machine learning,” International Journal of Pattern Recognition and Artificial Intelligence, vol. 36, no. 08, 2022. doi:10.1142/s021800142251003x

[9] S. K. Sharma, D. P. Sharma, and J. K. Verma, “Study on machine-learning algorithms in crop yield predictions specific to Indian agricultural contexts,” 2021 International Conference on Computational Performance Evaluation (ComPE), 2021. doi:10.1109/compe53109.2021.9752260   
[10] A. Puspaningrum, A. Sumarudin, and W. Putra, “Prediction of Weather Forecast for Smart Agriculture supported by Machine Learning,” 2022 5th International Conference of Computer and Informatics Engineering (IC2IE), 2022. doi:10.1109/ic2ie56416.2022.9970092

[11] A. Mahadware, A. Saigiridhari, A. Mishra, A. Tupe, and N. Marathe, “Rainfall prediction using different machine learning and deep learning algorithms,” 2022 2nd Asian Conference on Innovation in Technology (ASIANCON), 2022. doi:10.1109/asiancon55314.2022.9908857

[12] J.-L. Xu and Y.-L. Hsu, “Analysis of agricultural exports based on Deep Learning and Text Mining,” The Journal of Supercomputing, vol. 78, no. 8, pp. 10876–10892, 2022. doi:10.1007/s11227-021-04238-w

[13] R. Gamage et al., “Smart agriculture prediction system for vegetables grown in Sri Lanka,” 2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2021. doi:10.1109/iemcon53756.2021.9623259

[14] F. Raimundo, A. Gloria, and P. Sebastiao, “Prediction of weather forecast for smart agriculture supported by Machine Learning,” 2021 IEEE World AI IoT Congress (AIIoT), 2021. doi:10.1109/aiiot52608.2021.9454184

[15] K. Roy, S. S. Chaudhuri, and S. Pramanik, “Deep learning based real-time industrial framework for rotten and fresh fruit detection using semantic segmentation,” Microsystem Technologies, vol. 27, no. 9, pp. 3365–3375, 2020. doi:10.1007/s00542-020-05123-x

[16] S. Nosratabadi, S. Ardabili, Z. Lakner, C. Mako, and A. Mosavi, Prediction of food production using machine learning algorithms of multilayer perceptron and ANFIS, 2021. doi:10.21203/rs.3.rs-477719/v1