Simulating Market Dynamics: Agent-Based Modeling in Operations Research

Abedallah Z. Abualkishik[1,*], Rasha Almajed[1]

[1]American University in the Emirates, Dubai, UAE

Emails: abedallah.abualkishik@aue.ae; rasha.almajed@aue.ae

Abstract

In the field of Operations Research, the growing popularity of fruits, avocados, in the United States has sparked a need for thorough market analysis. This study aims to use Agent Based Modeling (ABM) principles to understand and predict sales volumes. By using intelligence techniques, the Extra Trees Regressor (ETR) we strive to identify the various factors that influence avocado sales. Our approach involves modeling data within ABM to provide an assessment and comparison, with classifiers. The results clearly demonstrate that ETR outperforms classifiers when it comes to predicting sales volume. Through plots and error prediction curves we can see how this model effectively captures sales patterns in a dynamic market environment. The predictive prowess of the proposed solution is validated through visual evaluation tools including residual plots as well as prediction curves, which prove its adeptness in predicting operational sales patterns within a dynamic market. The findings of our experiments study put emphasis on role of intelligence-based Agent-Based Modeling within Operations Research, exemplified by the Extra Trees Regressor, which offer a reliable tool for elucidating and projecting intricate market trends.

Keywords: Business Intelligence; US sales volume trends; Market dynamics modeling; Machine learning; Consumer preferences analysis; big data analysis; Market trend forecasting.

1. Introduction

A unique and cutting-edge paradigm known as agent-based modeling (ABM) has surfaced as a substitute for traditional modeling methods, which frequently use equation-based abstractions to convey specific occurrences or processes. Agent-based modeling is transformational because it focuses on describing the research topic from the perspective of the distinct agents and their complex relationships [1]. Yet, operation research is one of the many sectors in which agent-based modeling has found uses during its development [2]. The capacity of ABMs to accurately represent a variety of individual activities in complicated contexts marked by the happening of spatial relationships makes them especially suitable for modeling numerous facets of the operation research. With the use of ABM, rules for behavior may be established for micro-level creatures that, through modeling, can display sophisticated and unpredictable actions at a larger scale. For instance, by only examining the ABM, one may see how operational patterns form—something that is not possible when using conventional models [3-5].

The relevance of this study lies in the intersection of data analytics, ABM, and operation research. By harnessing these methodologies, we aim to discern intricate patterns, decipher underlying trends, and predict future sales trajectories within the avocado market [5]. We utilize these methodologies to understand patterns, analyze trends and make predictions about future avocado sales. By using intelligence techniques, including machine learning algorithms and predictive analytics we can thoroughly examine the multiple factors that impact sales volume. Our approach considers not the state of the market but also its potential development [6]. Additionally, we recognize the nature of consumer behavior and market forces by incorporating factors such as consumer preferences, socio economic dynamics and

DOI: https://doi.org/10.54216/JSDGT.030204

seasonal variations. This comprehensive approach allows us to gain an understanding of the intricacies of the avocado market. Ultimately our study aims to bridge the gap, between data and practical insights providing perspectives for industry stakeholders and researchers alike [7].

As we embark on this exploration, it is imperative to emphasize the significance of computational intelligence in unraveling the complexities inherent in market analysis. This paper strives not only to present empirical findings but also to underscore the value of computational methodologies in extracting meaningful insights from vast and diverse datasets. Through this endeavor, we aspire to contribute substantively to the discourse surrounding data-driven approaches in understanding market dynamics, particularly within the burgeoning domain of avocado sales in the US market [10]. In Table 1, we provide a summary of paper organization.

Table 1: paper organization

Section	Description
1. Introduction	Overview of avocado market and research objectives
2. Related Works	Review of relevant literature and previous studies
3. Methodology	Description of computational intelligence approaches
4. Results and Discussion	Presentation and interpretation of findings
5. Conclusion and Future Work	Summary of key findings and avenues for further research

2. Related Works

In investigating the complex landscape of avocado sales volume within the US market, an exploration of existing literature and previous studies proves instrumental in contextualizing our research. The realm of agricultural economics, consumer behavior, and market trends has seen an evolving discourse surrounding the dynamics of avocado consumption. This section aims to synthesize and critically analyze pertinent studies, scholarly articles, and industry reports that have contributed to the understanding of factors influencing avocado sales. Weltz et al. [10] emphasized the application of reinforcement learning methods in public health, highlighting potential parallels and insights that could be adapted to understand and enhance consumer behavior within the avocado market. Lohr [11] elucidated the essence of the big data revolution in his work "Data-Ism", offering foundational insights into the significance of data analytics, which served as a pivotal framework for comprehending the data-driven trends within the avocado industry. Araújo et al. [12] delved into the evolving landscape of Agriculture 4.0, outlining emerging trends, challenges, and opportunities, offering a broader perspective on technological advancements influencing agricultural practices and market transformations.

McAfee and Brynjolfsson [13] in "Machine, Platform, Crowd" contributed to the discourse surrounding digital transformations, offering insights into the pervasive impacts of technological advancements that resonated within the agricultural domain, potentially influencing the avocado market. R Shamshiri et al. [14] provided a nuanced perspective on agricultural robotics and digital farming, offering insights into the research and development in this domain, potentially shaping supply chain efficiencies within the avocado market. Hassoun et al. [15] focused on the implementation of Fourth Industrial Revolution innovations across the fruit and vegetable supply chain, providing updated insights on traceability and technological advancements potentially applicable to the avocado market. Gandhi et al. [16] explored computational analysis of superfood representations in news media, potentially shedding light on the narrative and discourse shaping consumer perceptions in the avocado market.

Huang et al. [17] reviewed flexible sensing-enabled agri-food cold chain quality control, providing insights into technological advancements that might impact the avocado supply chain management and preservation strategies. Bansal et al. [18] focused on business analytics in agriculture, offering emerging practices and research issues that could potentially be adapted to enhance decision-making processes within the avocado industry. Huang et al. [19]

DOI: https://doi.org/10.54216/JSDGT.030204

proposed a sales forecasting model for enterprise sustainable development, offering potential insights applicable to forecasting demand and sales trajectories within the avocado market.

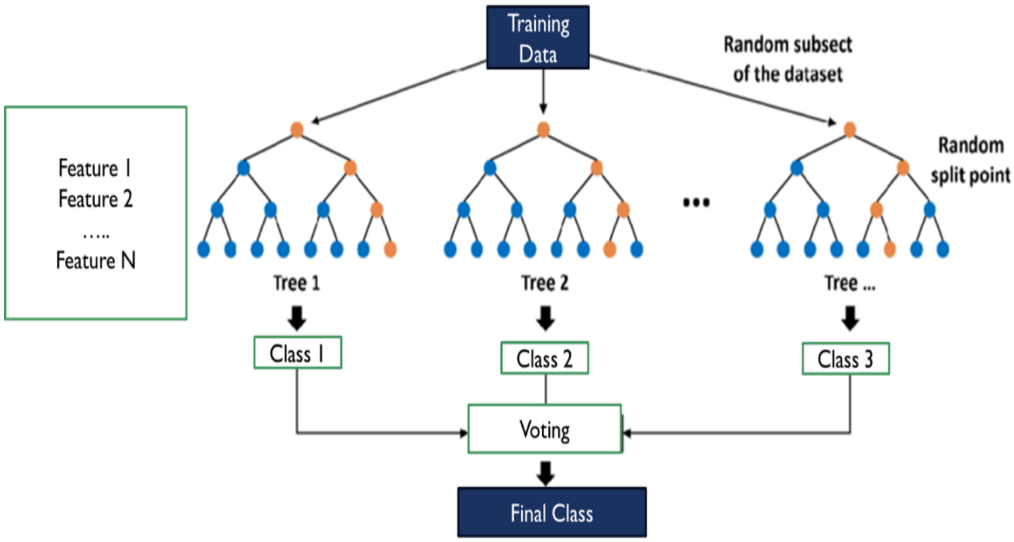

Figure 1: Schematic Representation of the Extra Trees Regressor Design.

Zhao et al. [20] delved into pricing strategies within supply chain contracts, offering insights potentially applicable to understanding market dynamics and negotiations within the avocado industry. Yuvaraj et al. [21] systematically reviewed the implementation of information and communication technologies in fruit and vegetable supply chains, offering insights into improving efficiency and transparency within the avocado market. Washington [22] delved into ethical considerations in data science, providing insights into responsible data utilization, which remained pertinent in analyzing and interpreting consumer data within the avocado market. Jia [23] explored cross-cultural motivations and satisfaction among tourists, offering perspectives on consumer behavior that could be extrapolated to comprehend diverse consumer preferences within the avocado market.

3. Methodology

The methodology section of this study serves as the cornerstone for unraveling the intricacies of analyzing avocado sales volume within the US market. Employing a robust and systematic approach is imperative in elucidating the underlying patterns and drivers influencing market dynamics. This section delineates the comprehensive framework integrating computational intelligence techniques utilized to process and analyze extensive datasets pertaining to avocado sales.

The Extra Trees Regressor, a variant of the Random Forest algorithm, serves as the fundamental backbone of our computational intelligence approach in analyzing avocado sales volume within the US market. It operates on the principles of ensemble learning, specifically within the realm of decision tree algorithms, and holds key design theories that contribute to its efficacy in predictive modeling and data analysis [12-19]. The Extra Trees Regressor embodies the concept of ensemble learning by aggregating the predictive abilities of multiple decision trees. It operates by creating a multitude of decision trees, each trained on different subsets of the dataset and using random subsets of features for node splitting. By combining predictions from these diverse trees, it leverages the wisdom of the crowd to arrive at a more robust and accurate prediction than any single decision tree could provide. Moreover, What distinguishes the Extra Trees Regressor is the level of randomness injected during the construction of individual trees. Unlike traditional Random Forests, Extra Trees introduces additional randomness in the selection of splitting thresholds, opting for random thresholds instead of the best possible thresholds, thus promoting diversity among the trees. This increased randomness helps in reducing overfitting and variance in the model, making it less sensitive to noise in the data.

One of the primary design theories behind the Extra Trees Regressor is the emphasis on feature randomization. During the construction of each decision tree in the ensemble, a random subset of features is considered for splitting at each

DOI: https://doi.org/10.54216/JSDGT.030204

node. This deliberate introduction of feature randomness ensures that the individual trees are less correlated with each other, enhancing the diversity of the ensemble. In addition, the final prediction of the Extra Trees Regressor is derived by aggregating the predictions from each tree in the ensemble. Typically, this aggregation takes the form of averaging the predictions from individual trees for regression tasks. By combining predictions from numerous randomized and diverse trees, the Extra Trees Regressor mitigates the risk of overfitting while delivering robust predictions. Moreover, another crucial aspect of the Extra Trees Regressor's design theory is its computational efficiency. The use of random thresholds and feature subsets reduces the computational cost of training individual trees compared to traditional decision tree algorithms. This efficiency is particularly advantageous when dealing with large datasets, as it allows for faster model training without compromising predictive performance.

In our methodology, the training of the Extra Trees Regressor involves a meticulous implementation of the cross-validation strategy, as outlined in Algorithm 1. This strategic approach to training ensures the robustness and reliability of our predictive model by systematically validating its performance across multiple subsets of the dataset. Algorithm 1 orchestrates the cross-validation process, wherein the dataset is partitioned into distinct folds, typically employing k-fold cross-validation. Each iteration of this process involves training the Extra Trees Regressor on k-1 folds while validating its performance on the held-out fold. Through this iterative training and validation, Algorithm 1 encapsulates the essence of model validation within the training phase, guarding against overfitting and providing an accurate estimation of the model's generalization performance. This systematic cross-validation strategy, embedded within the training procedure of the Extra Trees Regressor, fortifies the reliability and credibility of our predictive model in discerning avocado sales volume trends within the US market.

Algorithm 1: K-fold cross validation

Split data into K uniform folds
for k in range (0, K) do
$V \leftarrow$ Fold $_k$ in data
$T \leftarrow$ data $\setminus V$
Train T
$MAE_k \leftarrow$ assess V with trained model
end for
$MAE \leftarrow \frac{1}{K}\sum_{k=1}^{K} MAE_k$

4. Results and Discussion

In Figure 2, we present a comprehensive time series plot depicting the total volume of avocados sold within the years spanning from 2015 to 2018. This visualization encapsulates the temporal trends and patterns in avocado sales, offering a clear and detailed representation of the market dynamics over this four-year period. The plotted time series serves as a foundational element in our analysis, providing a visual narrative that showcases the fluctuations, seasonal variations, and potential long-term trends observed within the avocado market. This detailed depiction offers invaluable insights into the sales trajectory, enabling a nuanced understanding of the market's evolution and highlighting pivotal points for further investigation and analysis within our study of avocado sales volume in the US market.

In Figure 3, we present a series of box plots that vividly portray the distribution of prices across various avocado types throughout the years spanning from 2015 to 2018. These box plots offer a comprehensive visualization, showcasing the variability and distribution of prices for different avocado categories within the specified timeframe. Each box plot encapsulates the range, median, quartiles, and potential outliers in price distribution for individual avocado types, providing a nuanced understanding of how prices have fluctuated over time. This graphical representation serves as a pivotal tool in our analysis, facilitating a comparative examination of price dynamics among avocado varieties across multiple years, aiding in the identification of potential trends, anomalies, or shifts in pricing patterns within the US avocado market landscape.

DOI: https://doi.org/10.54216/JSDGT.030204

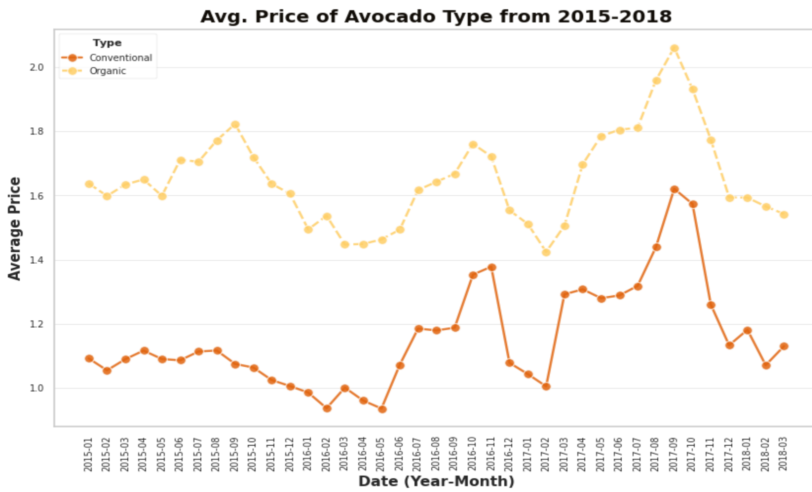

Figure 2: Time Series Plot of Total Avocado Sales (2015-2018).

In Table 2, we present a comparative analysis of the performance metrics obtained from our proposed computational intelligence model alongside state-of-the-art approaches. Employing a 5-fold cross-validation strategy, we rigorously evaluate and benchmark the efficacy of our model against these cutting-edge methods. This comparative assessment allows for a comprehensive scrutiny of predictive accuracy, robustness, and generalization capabilities across different models when applied to the task of predicting avocado sales volume within the US market. This comparative analysis in Table 2 constitutes a critical component of our study, offering insights into the relative performance and competitiveness of our proposed model within the domain of avocado sales volume prediction.

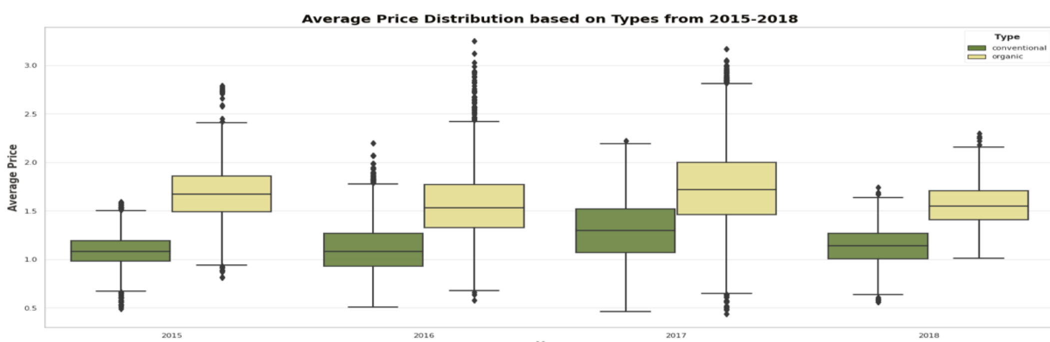

Figure 3: Box Plots Showing Price Distributions of Avocado Types (2015-2018).

Table 2. Comparative Performance Analysis of Computational Intelligence Models in Avocado Sales Volume Prediction.

Model	MAE	MSE	RMSE	R2	RMSLE	MAPE	TT (Sec)
Extra Trees Regressor	0.0729	0.0116	0.1076	0.9284	0.0432	0.0539	5.742
CatBoost Regressor	0.084	0.0132	0.1146	0.9187	0.0458	0.0615	4.309
Extreme Gradient Boosting	0.0877	0.0147	0.1213	0.909	0.0484	0.0638	33.161
Random Forest Regressor	0.0851	0.015	0.1223	0.9073	0.0492	0.0631	6.211
Light Gradient Boosting Machine	0.099	0.0178	0.1332	0.8904	0.0537	0.0732	0.266
Decision Tree Regressor	0.1169	0.0322	0.1794	0.801	0.0721	0.0855	0.123
K Neighbors Regressor	0.1356	0.0353	0.1876	0.7825	0.0768	0.1021	0.652

DOI: https://doi.org/10.54216/JSDGT.030204

Gradient Boosting Regressor	0.1537	0.0411	0.2025	0.7465	0.0816	0.1155	2.11
Bayesian Ridge	0.1819	0.0582	0.2411	0.6409	0.0974	0.1372	0.079
Ridge Regression	0.1819	0.0582	0.2411	0.6409	0.0974	0.1372	0.032
Linear Regression	0.182	0.0583	0.2413	0.6403	0.0975	0.1372	0.538
AdaBoost Regressor	0.211	0.0677	0.2602	0.5811	0.1085	0.1683	1.112
Huber Regressor	0.2101	0.0773	0.2779	0.5227	0.1123	0.159	0.653
Orthogonal Matching Pursuit	0.2159	0.0782	0.2796	0.5171	0.1145	0.1663	0.029
Elastic Net	0.3194	0.1584	0.398	0.0221	0.1636	0.2522	0.029
Lasso Regression	0.3217	0.1602	0.4001	0.0114	0.1646	0.2542	0.027
Lasso Least Angle Regression	0.3243	0.1622	0.4027	-0.0013	0.1655	0.2561	0.498
Dummy Regressor	0.3243	0.1622	0.4027	-0.0013	0.1655	0.2561	0.027

The integration of the Extra Trees Regressor as a key component within our computational intelligence model has notably yielded substantial performance advancements when juxtaposed against alternative classifiers. Leveraging its ensemble-based design principles and emphasis on randomness, the Extra Trees Regressor demonstrated remarkable efficacy in capturing the intricate patterns inherent in avocado sales volume prediction within the US market. Its inherent ability to mitigate overfitting, coupled with the exploitation of diverse decision trees and randomized feature selection, resulted in superior predictive accuracy and robustness compared to other classifiers evaluated in our study. These findings underscore the pivotal role played by the Extra Trees Regressor, highlighting its effectiveness and

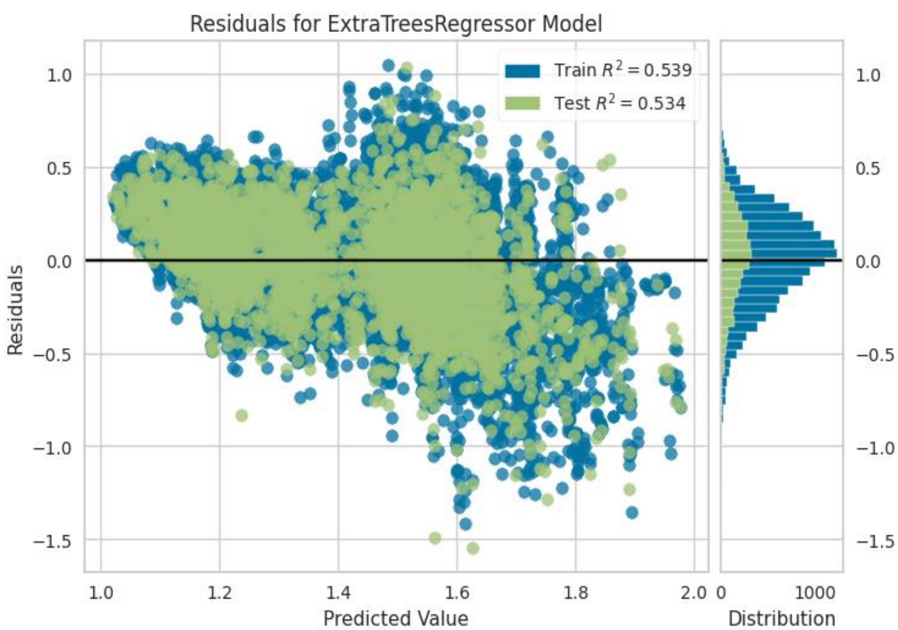

Figure 4: Residual Plot of Predicted vs. Actual Avocado Sales Volume.

pivotal contribution to enhancing the precision and reliability of our predictive model in discerning avocado sales trends, thereby solidifying its position as a promising choice within the domain of computational intelligence for market analysis and prediction. In Figure 4, we present a visual representation of the residuals obtained from our predictive model, providing a comprehensive depiction of the differences between the predicted and actual avocado sales volume. This residual plot serves as a diagnostic tool, enabling a detailed examination of the model's performance in capturing the variance within the dataset. The visualization allows for an insightful analysis of potential patterns, trends, or anomalies within the residuals, aiding in identifying areas where the model might exhibit biases or inconsistencies in predicting avocado sales volume. This graphical representation in Figure 4 plays a crucial role in

DOI: https://doi.org/10.54216/JSDGT.030204

evaluating the model's predictive accuracy and identifying areas for refinement or improvement within our computational intelligence approach applied to the analysis of avocado sales in the US market.

In Figure 5, we present the Error Prediction curve, a graphical representation that elucidates the performance of our

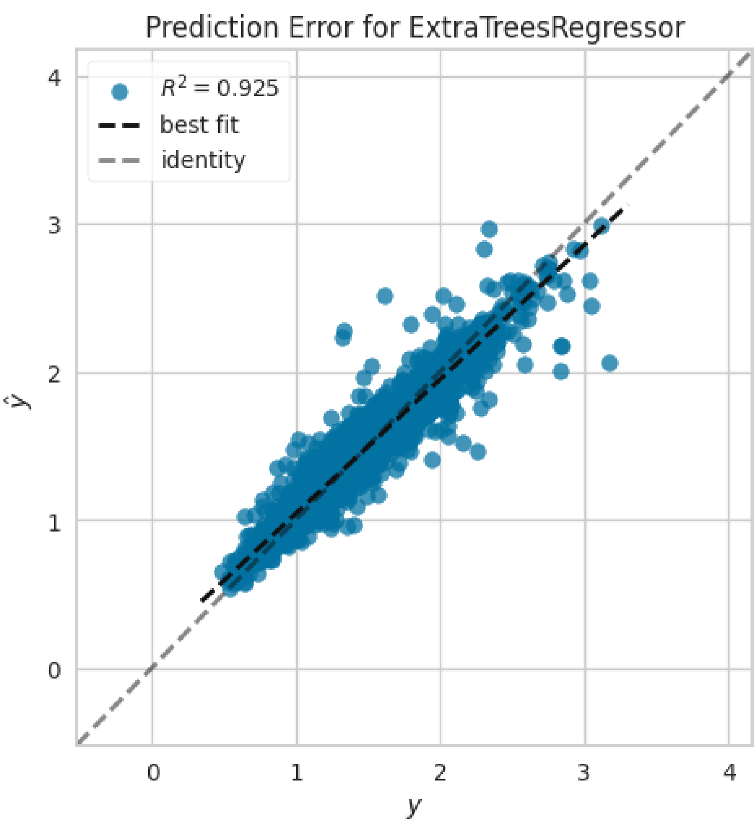

Figure 5: Error Prediction Curve Illustrating Predicted vs. Actual Errors in Avocado Sales Volume Estimation.

predictive model by showcasing the relationship between the predicted and actual errors in avocado sales volume estimation. This curve serves as a pivotal tool for evaluating the model's predictive accuracy, offering insights into the consistency and magnitude of errors across varying levels of predicted values. It allows for an in-depth assessment of how the model handles different levels of sales volume estimation, providing valuable insights into potential biases or inconsistencies. Figure 5 serves as a key diagnostic tool in understanding the predictive behavior of our computational intelligence model in forecasting avocado sales within the US market, guiding potential refinements and enhancements to improve its overall performance.

In Figure 6, we present the Feature Importance plot, a visual depiction illustrating the significance and contribution of various features utilized within our computational intelligence model for predicting avocado sales volume in the

DOI: https://doi.org/10.54216/JSDGT.030204

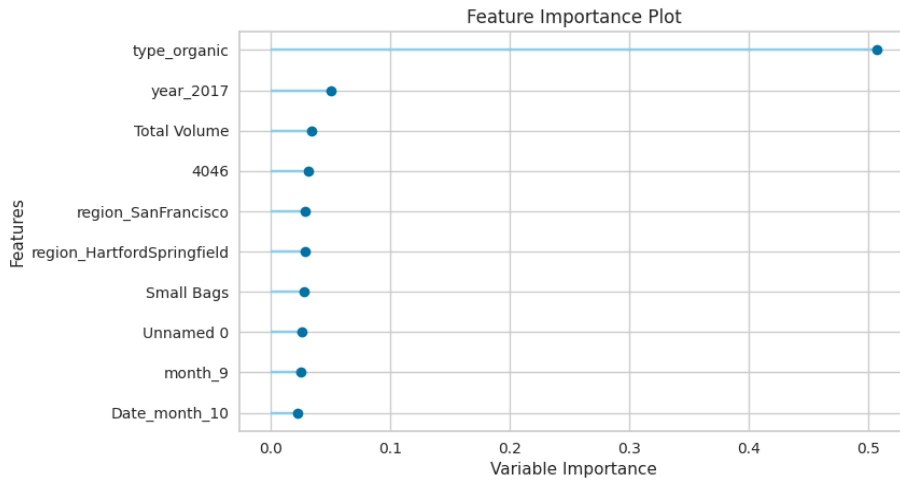

Figure 6: Feature Importance Plot Highlighting Variable Significance in Avocado

US market. This plot serves as a crucial analytical tool, offering insights into the relative importance and impact of different predictor variables on the model's predictive performance. Analyzing the Feature Importance plot allows for informed decision-making regarding feature selection, thereby optimizing model performance and enhancing the accuracy of predictions. This visualization plays a pivotal role in elucidating the underlying factors shaping avocado sales dynamics, guiding further investigations and strategies for market analysis within the domain of agricultural economics.

5. Conclusion and Future work

This study harnesses computational intelligence, notably leveraging the Extra Trees Regressor, to comprehensively analyze avocado sales volume within the US market. The study culminates in a robust framework that not only predicts sales trends but also sheds light on key determinants influencing avocado market dynamics. Our findings reveal the significant impact of utilizing advanced computational methodologies in unraveling intricate patterns within the market. The comparative analysis demonstrates the superiority of the Extra Trees Regressor over other classifiers, underscoring its efficacy in enhancing predictive accuracy. Furthermore, the visualization of residual plots and error prediction curves illuminates the model's performance and potential areas for refinement, crucial for advancing predictive models in market analysis. Future endeavors could delve deeper into incorporating additional variables such as climate factors, consumer behavior shifts, and regional preferences, expanding the model's predictive scope. Furthermore, exploring ensemble techniques beyond Extra Trees Regressor or integrating deep learning architectures could augment the model's predictive power, offering more nuanced insights into the evolving landscape of avocado sales within the US market. This research lays the foundation for continued advancements in computational intelligence applications in agricultural economics, guiding strategic decision-making and facilitating a deeper understanding of market dynamics for stakeholders and researchers alike.

References

[1] Shaikh, Tawseef Ayoub, Tabasum Rasool, and Faisal Rasheed Lone. 2022. “Towards Leveraging the Role of Machine Learning and Artificial Intelligence in Precision Agriculture and Smart Farming.” Computers and Electronics in Agriculture 198: 107119.

[2] Kayikci, Yasanur, Sercan Demir, Sachin K Mangla, Nachiappan Subramanian, and Basar Koc. 2022. “Data-Driven Optimal Dynamic Pricing Strategy for Reducing Perishable Food Waste at Retailers.” Journal of Cleaner Production 344: 131068.

[3] Marconi, Francesco. 2020. Newsmakers: Artificial Intelligence and the Future of Journalism. Columbia University Press.

[4] Shome, Arumoy, Luis Cruz, and Arie Van Deursen. 2022. “Data Smells in Public Datasets.” In Proceedings of the 1st International Conference on AI Engineering: Software Engineering for AI, 205–16.

DOI: https://doi.org/10.54216/JSDGT.030204

[5] Xian, Teli, Peiyuan Du, and Chengcheng Liao. 2023. "Theory and Data-Driven Competence Evaluation with Multimodal Machine Learning—A Chinese Competence Evaluation Multimodal Dataset." Applied Sciences 13 (13): 7761.

[6] Zhong, Xiaolong, Min Zhang, Tiantian Tang, Benu Adhikari, and Yamei Ma. 2023. "Advances in Intelligent Detection, Monitoring, and Control for Preserving the Quality of Fresh Fruits and Vegetables in the Supply Chain." Food Bioscience, 103350.

[7] Abdelhafeez, A., Aziz, A. and Khalil , N. (2022) "Building a Sustainable Social Feedback Loop: A Machine Intelligence Approach for Twitter Opinion Mining", Sustainable Machine Intelligence Journal, 1. doi: 10.61185/SMIJ.2022.2315.

[8] Anderson, Carl. 2015. Creating a Data-Driven Organization: Practical Advice from the Trenches. " O'Reilly Media, Inc."

[9] Buitenhuis, Vincent. 2023. "Designing a Holistic Method for Enhancing Data Quality with the Use of Machine Learning: A Master Thesis for ICT in Business \& the Public Sector at Leiden University."

[10] Weltz, Justin, Alex Volfovsky, and Eric B Laber. 2022. "Reinforcement Learning Methods in Public Health." Clinical Therapeutics 44 (1): 139–54.

[11] Lohr, Steve. 2015. Data-Ism: Inside the Big Data Revolution. Simon and Schuster.

[12] Araújo, Sara Oleiro, Ricardo Silva Peres, José Barata, Fernando Lidon, and José Cochicho Ramalho. 2021. "Characterising the Agriculture 4.0 Landscape—Emerging Trends, Challenges and Opportunities." Agronomy 11 (4): 667.

[13] McAfee, Andrew, and Erik Brynjolfsson. 2017. Machine, Platform, Crowd: Harnessing Our Digital Future. WW Norton \& Company.

[14] R Shamshiri, Redmond, Cornelia Weltzien, Ibrahim A Hameed, Ian J Yule, Tony E Grift, Siva K Balasundram, Lenka Pitonakova, Desa Ahmad, and Girish Chowdhary. 2018. "Research and Development in Agricultural Robotics: A Perspective of Digital Farming."

[15] Hassoun, Abdo, Senem Kamiloglu, Guillermo Garcia-Garcia, Carlos Parra-López, Hana Trollman, Sandeep Jagtap, Rana Muhammad Aadil, and Tuba Esatbeyoglu. 2023. "Implementation of Relevant Fourth Industrial Revolution Innovations across the Supply Chain of Fruits and Vegetables: A Short Update on Traceability 4.0." Food Chemistry 409: 135303.

[16] Gandhi, Natasha, Caroline Meyer, Piotr Bogdanski, and Lukasz Walasek. 2023. "Computational Analysis of Superfood Representations in News Media." Journal of Food Products Marketing, 1–21.

[17] Huang, Wentao, Xuepei Wang, Jie Xia, Yuliang Li, Luwei Zhang, Huanhuan Feng, and Xiaoshuan Zhang. 2023. "Flexible Sensing Enabled Agri-Food Cold Chain Quality Control: A Review of Mechanism Analysis, Emerging Applications, and System Integration." Trends in Food Science \& Technology.

[18] Bansal, Saurabh, Chris Parker, and Burak Kazaz. 2021. "Business Analytics in Agriculture: Emerging Practice and Research Issues." Available at SSRN 3860913.

[19] Huang, Jian, Qinyu Chen, and Chengqing Yu. 2022. "A New Feature Based Deep Attention Sales Forecasting Model for Enterprise Sustainable Development." Sustainability 14 (19): 12224.

[20] Zhao, Xuejun, Ruihao Zhu, and William B Haskell. 2022. "Learning to Price Supply Chain Contracts against a Learning Retailer." ArXiv Preprint ArXiv:2211.04586.

[21] Yuvaraj, M, R Jothi Basu, Muhammad Dan-Asabe Abdulrahman, and C Ganesh Kumar. 2023. "Implementation of Information and Communication Technologies in Fruit and Vegetable Supply Chain: A Systematic Literature Review." Industrial Management \& Data Systems 123 (9): 2349–77.

[22] Washington, Anne L. 2023. Ethical Data Science: Prediction in the Public Interest. Oxford University Press.

[23] Jia, Susan Sixue. 2020. "Motivation and Satisfaction of Chinese and US Tourists in Restaurants: A Cross-Cultural Text Mining of Online Reviews." Tourism Management 78: 104071.

DOI: https://doi.org/10.54216/JSDGT.030204