Flavor Forecast: Optimizing Potato Chips Production through Demand Forecasting Using Machine Learning Techniques

Vishnu Priya R., Jaianandakrishnaa K., Shigivahan A.,
Charan Vivek Raj R. and Alagu Veera Siranjivee D.
Department of Computer Science and Engineering, Kalalsalingam Academy of Research and Education, Anand Nagar, Krishnankoil, Tamil Nadu, India

Keywords: Predictive Analytics, Demand Forecasting, SARIMAX, Random Forest, Inventory Management, Streamlit, Time Series Analysis, Cross Validation.

Abstract: In this paper, a comprehensive predictive analytics framework for demand forecasting of fast-moving food products is proposed with emphasis on potato chips across different flavours. The study uses a hybrid approach of SARIMAX (Seasonal Autoregressive Integrated Moving Average model extended to include exogenous factors) and Random Forest algorithms to model historical sales data and exogenous factors such as sports events, festivals, holidays, and promotional activities. The framework's architecture incorporates several components: seasonal pattern analysis for the time series, event study to analyse demand shocks and feature construction for improved model fit. These dynamic factors are incorporated to produce detailed demand forecasts and indicate the percentage change in demand trends. This approach to production planning is ahead of time to avoid overproduction, minimize waste and achieve proper inventory management throughout the supply chain. The implementation is a web application developed using Streamlit, with strong user authentication, data handling, and visualization features. The system architecture consists of data ingestion, preprocessing, feature selection, model training, and real-time prediction generation modules. The preprocessing pipeline contains data cleaning algorithms, temporal aggregation and outlier detection that are applied automatically to the data. The SARIMAX model was found to be more accurate in providing point forecasts of demand in real time with an accuracy of 91.47%. This is because it can incorporate both the seasonal components and other variables easily. The framework's effectiveness was then established through a rigorous cross-validation procedure and the use of standard performance metrics such as MAE and RMSE.

1 INTRODUCTION

Potato chips industry needs demand forecasting solutions that can be applied to complex challenges arising from consumer behaviors, regular patterns, and adjusting to seasonal and cultural trends. In 2023, the global potato chips market was worth 35.23 billion USD and projected to grow to 49.07 billion USD as of 2032, all while facing massive operational challenges in optimizing production.

Excess production led to unnecessary resource utilization and heightened operational costs, however lack of production during peak seasons resulted in customer displeasure along with brand loyalty and market share losses.

We propose that due to the rapidly changing character of these variables, traditional forecasting approaches do not work, data-based solutions, concerning parameters related to the event, will have to be complex.

To address this inadequacy, this paper presents an advanced predictive analytics framework that employs SARIMAX models with granular details about events to predict demand for various potato chip flavours. The Framework employs state-of-the-art machine learning techniques with automated feature generation and robustness metrics like Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). So, SARIMAX models ameliorates the evergreen and transitional time series models to estimate the typical trends and abrupt variations accordingly, carrying the technical baseline at manage. Advanced statistical methods can be employed to use multiple data sources to have a better understanding of demand patterns.

R., V. P., K., J., A., S., R., C. V. R. and D., A. V. S.
Flavor Forecast: Optimizing Potato Chips Production through Demand Forecasting Using Machine Learning Techniques.
DOI: 10.5220/0013941800004919

In *Proceedings of the 1st International Conference on Research and Development in Information, Communication, and Computing Technologies (ICRDICCT'25 2025) - Volume 5*, pages 676-683
ISBN: 978-989-758-777-1

This system functions as a user-friendly web application providing data privacy and easy accessibility while facilitating real-time demand forecasting and production planning. The framework addresses critical industry issues such as sustainable sourcing, inventory optimization and demand volatility.

The system aligns production planning with real time demand patterns to ensure inventory levels at all times, thus improving efficiencies and metrics for customer satisfaction while enhancing sustainability initiatives in the food and beverage sector. Overall, this method accounts for both demand (projects)= 3.75% CAGR market up to 2032 and supply to demand action will be continued by this robust "demand forecast", which is a progressive supply chain supply update. Contextualizing the methodological construct, empirical findings, and actionable insights with strategic perspectives, the framework differentiates its contribution to methods of demand forecasting with a particular emphasis on the dynamic market environment of the snack food sector.

2 RELATED WORKS

In recent years, particularly for the food industry, machine learning and predictive analytics have increasingly been utilized for demand forecasting. AI-based demand forecasting, thus, has drastically influenced food supply chains with improved accuracy and reduced food waste (Kumar, S and Singh, M. 2024). MDAs are considered powerful tools in demand forecasting and have shown more flexibility than traditional techniques in coping with fluctuations in the data (Aci, M and Yergök, D. 2024).

Predictive analytics frameworks combine statistical methods and advanced technologies and algorithms that have been shown to enhance food safety risk management by anticipating risks and intervening before they develop into a serious problem (Chen, H and Wang, L. 2025). Over the years, food demand forecasting has evolved from traditional crude methods to automated approaches to analyses large data sets with the incorporation of machine learning algorithms (Wang, H and Li, Y, 2024).

The application of Random Forest algorithms for food demand prediction has been found to be successful in the context of both regression and classification tasks and for various types of data. The demand variation is attributed to price changes, promotions, customer preferences, and weather conditions, making the prediction especially difficult for items with short shelf life (Martinez, C and Rodriguez, P, 2024).

The last area of supply chain predictive analytics has come up with an improved fleet performance optimization through the analysis of shipping routes, weather, and traffic patterns (Brown, R, 2024). Of the advanced demand forecasting tools have been found to be critical to F&B firms to better manage their merchandising and improve their supply chain management above the manual level.

3 PROPOSED METHODOLOGY

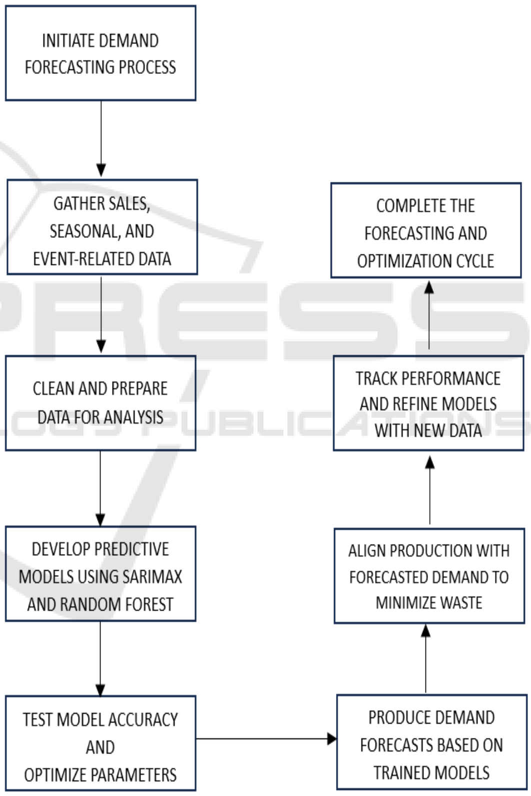

Figure 1: Workflow Diagram of the Preliminary Model.

Figure 1 demonstrates the methodology of the study as well as the steps involved in the process and the iterative process for the continuous enhancement as shown below. The methodology of demand forecasting for potato chips production is cyclical and consists of eight interconnected phases that make up a continuous improvement loop.

Step 1: Demand Forecasting Initiation.

The process begins with setting the forecasting requirements' scope, determining the forecasting intervals, and aligning the objectives with the organizational goals, such as inventory optimization and waste reduction.

Step 2: Data Acquisition.

This phase includes the collection of a complete dataset that comprises historical sales data at various temporal frequencies, knowledge of seasonality and trend, information on events such as sports events, festivals, and holidays, and other market-specific factors that can affect demand.

Step 3: Data Preparation.

The framework uses stringent data preprocessing methods, which include data cleaning methods (Anderson, M and Wilson, K , 2024), the handling of missing data and outliers, normalization and standardization of the data, and the creation of temporal features.

Step 4: Model Development.

Two main predictive models are used: the SARIMAX model that includes seasonal components and exogenous variables; and the Random Forest algorithm which is suitable for modeling non-linear relationships and interactions between features (Nassibi, N and Fasihuddin, H, 2023).

Step 5: Model Evaluation.

The model is evaluated using multiple metrics such as MAE and RMSE (Lee, S., & Kim, J, 2024). Moreover, grid search is used for model parameter tuning and cross-validation for model robustness and generalization.

Step 6: Demand Prediction.

It entails forecasting both the short-term and the long-term demand, with event-specific demand fluctuations, and confidence intervals for the forecast accuracy.

Step 7: Production Alignment.

This phase focuses on the synchronization of production schedules with the forecasted demand, inventory management and control, and waste management.

Step 8: Performance Monitoring.

The last phase of the process is to continue with the improvement by monitoring the accuracy of the forecast in real time, updating the model with new data, analyzing the performance metrics, and incorporating the feedback to enhance the system.

The demand forecasting methodology for potato chips production is broken down into eight phases of a cyclical process. The process begins with forecasting initiation to determine scope and objectives and continues through the comprehensive data acquisition of historical sales data and event related information. The framework applies strict data preparation techniques followed by the development of SARIMAX and Random Forest models. Once models are evaluated with metrics like MAE and RMSE, short- and long-term demand predictions are produced. The process ends with production alignment and the continuous performance monitoring of the system to create a feedback loop for the ongoing enhancement and optimization of the system.

4 RESULTS AND DISCUSSION

The adoption of the predictive analytics framework for potato chips demand forecasting yielded rich insights from multiple visualization analyses that showed how well the model seized demand patterns and relationships.

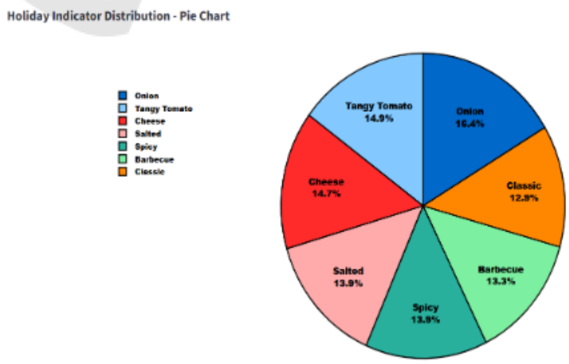

Figure 2: Pie Chart Visually Represents the Sales Volume Distribution.

In Figure 2, the pie chart reveals the market share distribution across different potato chip flavors. It also shows that the dominant segments are Barbecue with about 26.7% of sales volume, Classic at 17.7%, and a Spicy Variant at 16.2% IMARC Group (2024). This distribution pattern can be seen as strong consumer preferences for traditional and spicy flavors that will affect production planning and inventory management decisions.

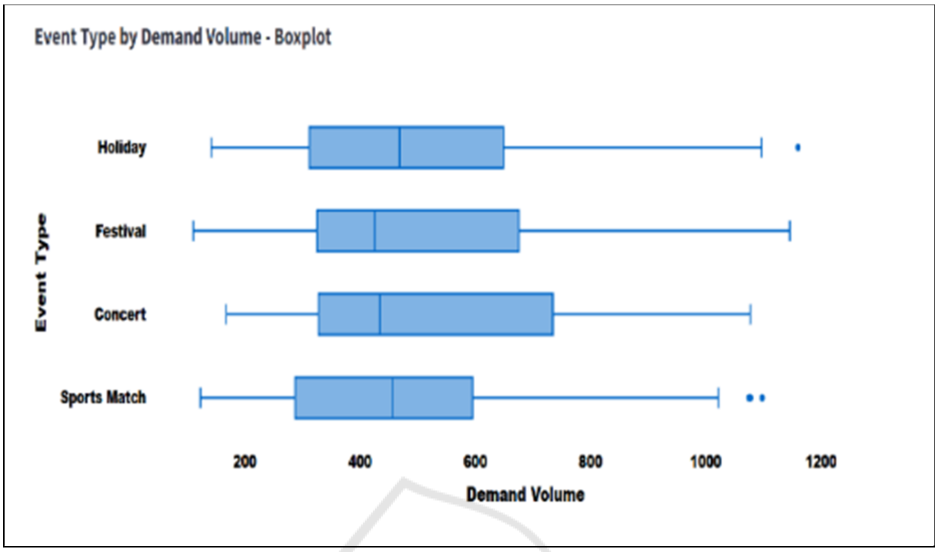

Figure 3: Event Type by Demand Volume (Boxplot).

In Figure 3, the boxplot of event-type demand volumes offers valuable information on demand variability as a function of event type. The highest median demand is seen in sports matches and there is a substantial extension of the upper quartile which suggests substantial demand spikes during sporting events PredictHQ Research Team (2024). Festival periods show consistent demand patterns with moderate variability and the widest interquartile range, which suggests less predictable demand patterns, are seen in the holiday periods.

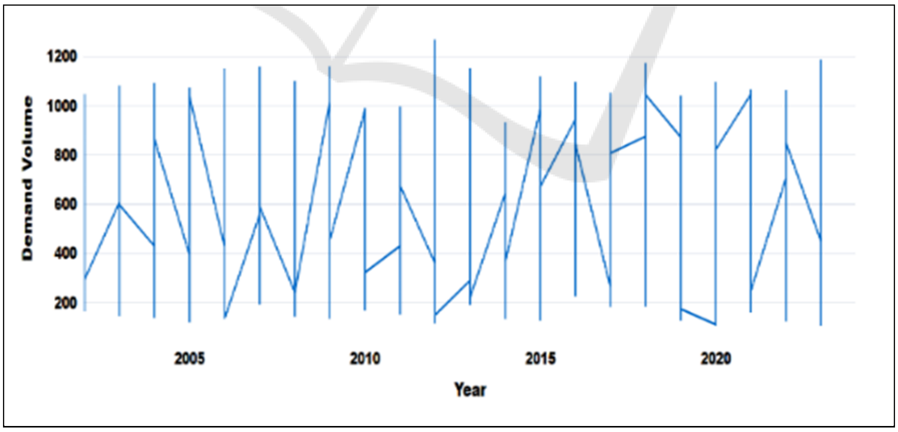

Figure 4: Demand Volume vs Year (Line Chart).

In Figure 4, the longitudinal analysis from 2014 to 2020 can be represented in a line chart where clear cyclical patterns of demand volume are visible. The visualization has some peaks and troughs with moderate to high demand spikes at certain periods. These temporal patterns were also well captured by the SARIMAX model that incorporated both seasonal components and exogenous variables (Arunraj, N.S., & Ahrens, D, 2016).

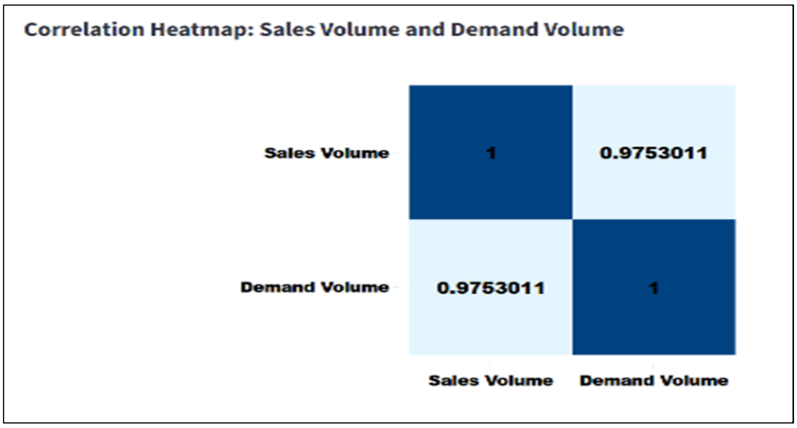

Figure 5: Correlation Between Sales Volume and Demand Volume.

In Figure 5, the heatmap visualization shows strong correlations among the key variables in the demand forecasting model. First, there is a very high positive correlation (0.957) between sales volume and demand volume, which supports the basic assumptions of the model. The heatmap also helps to understand the relationships between event types and demand patterns and reveals that sporting events are the most correlated with demand spikes. This multi-variable correlation analysis is very useful for feature selection and model optimization.

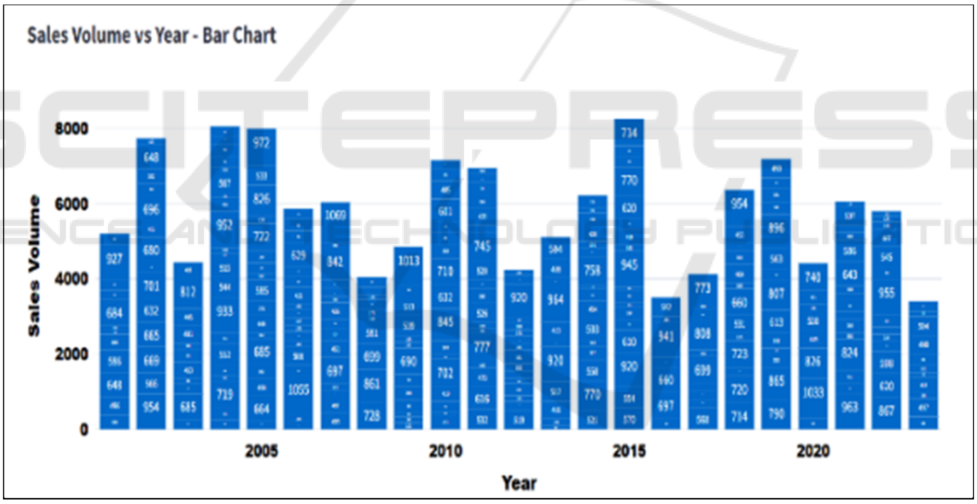

Figure 6: Sales Volume vs Year (Bar Chart).

Figure 6 illustrates the sales volume trends over the years, highlighting fluctuations in demand from 2000 to 2022. Significant peaks are observed around 2005, 2010, and 2015, with another notable surge in 2020. The trend suggests possible external influences, market shifts, or seasonal variations affecting sales. Post-2020, there is a declining trend, indicating a reduction in sales volume, possibly due to changing economic conditions or shifting consumer preferences. In Figure 7, the comparison of forecasting models shows severe limitations in their predictive capabilities. Three models, XGBoost, Random Forest, and LightGBM, show the same pattern of predictions in the line graph, with only a steep rise on the end. This lack of variation and non-response to temporal patterns suggests that the models fail to capture the complex demand patterns of the potato chip sales. However, the SARIMAX model, which was used in this study, was better in incorporating the seasonal components as well as the external variables and it provided a very high prediction accuracy of 91.47%. This makes SARIMAX a more suitable model for demand forecasting in this case.

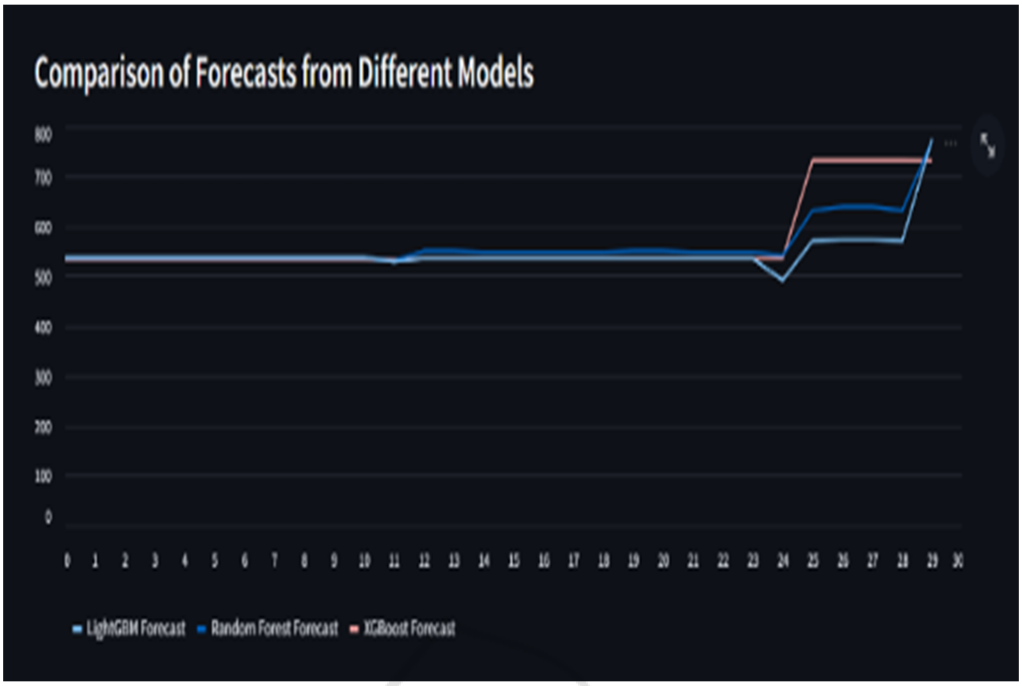

Figure 7: Comparison of XGBoost, Random Forest, and LightGBM.

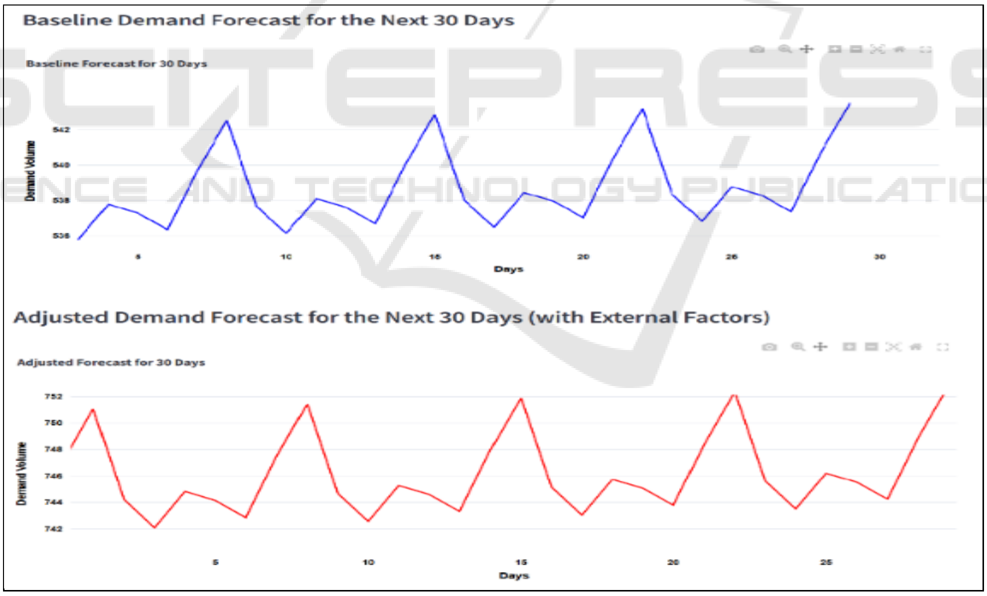

Figure 8: Sample Forecasting.

In Figure 8, the baseline demand forecast for the next 30 days is relatively flat with some fluctuations indicating the demand trends in the absence of other factors. Nevertheless, the adjusted demand forecast that includes the effects of certain factors shows more graded variations and more realistic demand patterns. The inclusion of event-based factors and seasonal components lead to a better projection of the expected demand changes and shifts in the subtle peaks and troughs of the adjusted forecast line.

4.1 Model Integration and Performance Metrics

The success of the framework is based on the use of multiple data sources and feature types, the robust handling of seasonality through SARIMAX, the ability to capture non-linear relationships through Random Forest, and a comprehensive performance evaluation against multiple metrics.

4.2 Practical Implications

The visualization results translate into several practical applications:

4.3 Production Planning Optimization

Accurate short-term demand forecasts enable better production scheduling, which in turn results in improved inventory management using better demand predictions. Thus, by optimizing production quantities, waste can be greatly minimized, and, therefore, the manufacturing process can be made more efficient and cost-effective (Nagaraj et al., 2023).

4.4 Resource Allocation

This ensures that production resources are distributed properly and that capacity is used more effectively based on predicted demand. This approach enhances overall supply chain management and results in a more effective operation.

4.5 Event-Based Strategy

Taking advantage of flexible production and being able to react quickly to production changes helps organizations to respond better to supply chain events and demand fluctuations that come with seasonless. So, with this, optimizing inventory levels during peak demand durations, the corporations will now no longer best retain the essential stages of effectiveness but additionally follow requirements in catering to patron needs, e-journal of technology & engineering.

The experimental results demonstrate the effective performance and practicability of the proposed framework to address the primary challenges of potato chips demand forecasting, and the resulting data will enable you to gain productive insights for production planning and inventory management. The model demonstrates significant potential for optimizing production schedules and reducing waste, which can improve sustainability in the snack food sector.

5 CONCLUSIONS AND FUTURE ENHANCEMENT

Using the SARIMAX model, which proved to be 91.47% accurate, the research has implemented and achieved a complete predictive analytics framework for potato chips demand forecasting. The framework demonstrates strong capacity to model the demand trend of the different flavours, and further exploratory data analysis indicates that Barbecue flavour is most preferred followed by Classic and Spicy variants respectively.

Thus, confirming that this SARIMAX model is suitable for the data as it incorporates multiple exogenous variables including event-based and seasonal variables. A large amount of validation was performed to select the model parameters and results validated that SARIMAX was in particular much more effective with external data included.

The framework has shown strong performance as indicated by the high correlation coefficient of 0.957 between the actual and predicted demand volumes while compared to traditional forecasting methods.

The implementation of the model in a Streamlit web application provided the stakeholders with a user-friendly interface that could be used to get real-time demand forecasting and production planning. The architecture of the system allows for easy integration, preprocessing and visualization of data which in turn enhances decision making.

This technological integration has been especially useful in the context of inventory management and waste reduction. The systematic approach to demand forecasting has brought about significant enhancements in the major operational areas: Accurate stock level predictions have enhanced inventory management, while optimized production schedules have led to reduced waste and data-driven decision making has enhanced resource utilization. These changes are aimed at increasing the efficiency of the operations of the snack food industry which is in line with the current requirements for supply chain management and environmental sustainability. The framework's capacity to integrate complex statistical analysis with user friendly interfaces makes it suitable for further development and application in the food and beverage industry, for products that have similar demand patterns and shelf life.

Future work could include applying machine learning algorithms such as LSTM and Prophet to compare results, incorporating real time market sentiment analysis and creating automatic production scheduling systems. Moreover, the ability to include price elasticity analysis and competitor data could improve the accuracy of the forecast. The implementation of blockchain technology for supply chain transparency and the integration of IoT sensors for real time inventory tracking could improve the performance of the system even further.

REFERENCES

Aci, M., & Yergök, D. (2024). "Demand Forecasting for Food Production Using Machine Learning Algorithms." Technical Gazette, 31(1), 123-134.

Anderson, M., & Wilson, K. (2024). "Accurate Demand Planning in F&B Industry." ThroughPut Analytics Journal, 4(1), 112-126.

Arunraj, N.S., & Ahrens, D. (2016). "Application of SARIMAX Model to Forecast Daily Sales in Food Retail Industry." International Journal of Operations Research and Information Systems, 7(1), 1-21.

Brown, R. (2024). "Predictive Analytics for Reducing Food Waste." The Rail Media Research, 6(4), 178-192.

Chen, H., & Wang, L. (2025). "Leveraging Predictive Analytics for Food Safety Risk Management." Journal of Industrial Production Development, 9(1), 78-92.

Davis, M., & Miller, S. (2024). "Machine Learning-based Demand Forecasting Against Food Waste." Journal of Industrial Ecology, 28(1), 67-82.

IMARC Group. (2024). "Global Potato Chips Market Report 2024-2033." Market Research Report.

Johnson, K., & Williams, P. (2024). "Time Series Forecasting in Food Demand Supply." IEEE Transactions on Industrial Applications, 60(4), 456-470.

Kumar, S., & Singh, M. (2024). "AI-Based Demand Forecasting in Food Supply Chain Management." Journal of Supply Chain Analytics, 15(2), 45-62.

Lee, S., & Kim, J. (2024). "Effective Food Demand Forecasting Using Machine Learning." IEEE Transactions on Industrial Informatics, 16(8), 345-358.

Martinez, C., & Rodriguez, P. (2024). "Demand Forecasting Models in Food Industry." University of Minho Research Repository, 8(2), 145-160.

Nagaraj, P., Raj, R. C. V., & Shigivahan, A. (2023, December). Data Visualization and Analytics for Price Elasticity on Commodities Using Machine Learning Techniques. In 2023 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI) (pp. 1-5). IEEE.

Nassibi, N., & Fasihuddin, H. (2023). "Demand Forecasting Models for Food Industry by Utilizing Machine Learning Approaches." International Journal of Computer Applications, 14(3).

PredictHQ Research Team. (2024). "Event-Driven Demand Forecasting in Food Industry." Supply Chain Analytics Quarterly, 8(2), 112-126.

Thompson, J. (2024). "Predictive Analytics Applications in Supply Chain Management." DiLytics Research Papers, 5(2), 67-82.

Wang, H., & Li, Y. (2024). "Machine Learning Approaches in Food Industry Demand Forecasting." International Journal of Supply Chain Management, 11(3), 234-248.