IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, March 2026
Impact Factor: 8.2

Food Court Sales Analysis and Prediction Using Machine Learning

Shravani Shankar Chakor[1], Lawanya Sandeep More[2],
Tanuja Santosh More[3], Gitesh Bharat Khairnar[4], Mr. A. S. Dhomase[5]
Students, Department of Computer Engineering[1-4]
Guide, Department of Computer Engineering[5]
Matoshri Aasarabai Institute of Technology and Research Centre, Eklahare, Nashik

Abstract: *Food sales prediction is a time series forecasting task that plays a crucial role in minimizing food wastage and improving business efficiency. This paper presents a comprehensive system for food court sales analysis and demand prediction using machine learning techniques. The proposed system collects transactional data from multiple food stalls and applies statistical and machine learning models including ARIMA, SARIMA, Triple Exponential Smoothing, Support Vector Machines (SVM), and Prophet for accurate sales forecasting. The system incorporates a web-based dashboard that visualizes stall-wise sales, revenue by category, hourly trends, and recent transactions. A Stacked SVM model demonstrated superior prediction accuracy over other methods when evaluated using RMSE and MAPE metrics. The system assists food court managers in inventory management, demand forecasting, staffing optimization, and strategic decision-making.*

Keywords: Machine Learning, Sales Prediction, Food Court, ARIMA, SVM, Time Series Forecasting, Dashboard, Data Analytics

I. INTRODUCTION

Forecasting is the process of predicting future events or trends based on historical data and current conditions. It is a vital tool in domains such as retail, finance, weather prediction, and food services. In the food industry, accurate short-term sales forecasting enables businesses to minimize expired products, avoid stock-outs, and optimize resource allocation.

Food courts in malls, universities, airports, and commercial complexes generate large volumes of transactional data daily. Efficient analysis of this data is critical for improving operational efficiency, demand forecasting, and customer satisfaction. Traditional statistical methods such as Moving Average, Exponential Smoothing, and ARIMA have been widely used; however, they often fail to capture complex, nonlinear patterns present in real-world data.

This paper proposes a machine learning–based system for food court sales analysis and prediction. The system processes structured transaction data from multiple food stalls including attributes such as date, time, stall name, item category, quantity sold, and payment type. It applies multiple forecasting models and presents insights through interactive web-based dashboards.

Several different forecasting methods are employed in this work: (1) Trend Analysis — examining past data to identify patterns; (2) Regression Analysis — analyzing relationships between variables; (3) Time Series Analysis — ARIMA, SARIMA, and Exponential Smoothing; and (4) Machine Learning Models — SVM, Random Forest, and Stacked SVM.

II. LITERATURE SURVEY

The rapid growth of the food service industry has led to the generation of large volumes of transactional and customer-related data. Traditional statistical methods have been widely used for sales analysis; however, machine learning (ML) techniques have emerged as powerful tools for food court sales analysis and prediction.

DOI: 10.48175/IJARSCT-32143

Traditional Sales Forecasting Approaches

Early studies on sales analysis in the food and retail sectors relied heavily on statistical and econometric techniques. Time series models such as Moving Average, Exponential Smoothing, and Autoregressive Integrated Moving Average (ARIMA) were commonly used for predicting sales trends. These models assume linearity and stationarity in data and perform reasonably well for short-term forecasting with stable demand patterns. Several researchers applied regression-based models to analyze the relationship between sales and explanatory variables such as price, promotions, and customer traffic.

Machine Learning in Food Sales Analysis

With advancements in data collection technologies such as Point of Sale (POS) systems and digital payment platforms, machine learning has become increasingly popular in food sales analysis. Supervised learning algorithms such as Linear Regression, Decision Trees, Support Vector Machines (SVM), and k-Nearest Neighbors (k-NN) have been widely explored for predicting food sales. Studies have shown that Decision Tree-based models outperform traditional statistical techniques due to their ability to model nonlinear relationships. Random Forest and Gradient Boosting methods further improve prediction accuracy by combining multiple weak learners.

Deep Learning and Time-Series Forecasting

Recent research highlights the growing use of deep learning models for sales prediction in the food industry. Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, are widely used for time-series forecasting due to their ability to retain long-term dependencies. Several studies applying LSTM models to restaurant and food sales data report improved forecasting accuracy, especially for daily and hourly sales prediction.

Feature Engineering and Influencing Factors

Feature selection and engineering play a crucial role in food court sales prediction. Commonly used features include transaction timestamp, vendor ID, item category, price, discounts, and payment method. Temporal features such as weekday/weekend, festival days, and academic calendars significantly influence sales patterns. Recent literature emphasizes the inclusion of contextual data such as weather information, footfall counts, and online reviews to enhance prediction accuracy.

III. REQUIREMENTS ANALYSIS

3.1 Stakeholder Identification

Primary Stakeholders include Food Court Owners/Management (use sales analysis and prediction reports for decision-making), Food Stall/Outlet Vendors (analyze individual stall performance), System Administrators (manage user access and system maintenance), and Business Analysts (interpret sales trends and provide recommendations).

Secondary Stakeholders include Data Scientists/ML Engineers (develop and optimize prediction models), Web Developers (design and maintain the application), Inventory and Supply Chain Managers (plan stock procurement), and the Marketing Team (design promotional campaigns).

Tertiary Stakeholders include Customers (benefit from improved service), Suppliers/Vendors (adjust supply quantities based on predicted demand), and Investors/Business Partners (evaluate business performance).

3.2 Functional Requirements

1. User Authentication and Authorization: Secure login with role-based access for Administrator, Manager, and Vendor roles.

2. User Role Management: Different permissions per role — admins manage users and data, managers view overall performance, vendors access only their stall information.

3. Sales Data Entry and Storage: Structured collection and storage of date, time, stall name, food item, quantity sold, and total price with both manual entry and bulk upload support.

4. Data Preprocessing and Cleaning: Handling missing values, removing duplicate entries, correcting inconsistencies, and formatting data before analysis.

5. Historical Sales Analysis: Daily, weekly, and monthly sales summaries with stall and item comparisons.
6. Sales Trend Identification: Detection of peak sales hours, high-demand days, seasonal variations, and recurring purchase patterns.
7. Data Visualization: Graphs, charts, and dashboards displaying sales trends, comparisons, and forecasts.
8. Machine Learning Model Training: Training on historical data using regression and time-series algorithms with performance evaluation using RMSE and MAPE.
9. Sales Prediction and Forecasting: Future sales prediction for specific time periods, food items, or stalls.
10. Inventory Planning Assistance: Forecast-based assistance for raw material purchases and food preparation quantities.
11. Report Generation: Sales summaries, trend analysis, and forecasts in downloadable PDF or Excel formats.
12. Data Security and Backup: Protection of sensitive data and regular backups to prevent loss.

3.3 Non-Functional Requirements

- Performance: fast page loading, data retrieval, and prediction generation.
- Scalability: stable performance as data volumes and users grow.
- Reliability: consistent, accurate results with quick recovery from failures.
- Security: authentication, authorization, secure storage, and firewall protection.
- Usability: intuitive interface with minimal training required.
- Compatibility: support for Chrome, Firefox, Edge on desktops, laptops, and tablets.
- Maintainability: modular design enabling easy updates and bug fixes.
- Backup and Recovery: regular backups with fast data restoration capability.

IV. SYSTEM DESIGN

4.1 Architectural Design

The system follows a multi-tier architecture. Order data from multiple food stalls is transmitted over the internet through a firewall to a web server. The web server enqueues requests for parallel processing by the sales analysis and prediction modules. Results are stored in a centralized database for reporting and decision-making.
The key system components are: (1) User Interface — responsive web frontend for login, data entry, and dashboards; (2) Authentication Module — role-based access control; (3) Data Collection Module — manual entry and bulk CSV/Excel upload; (4) Database Management System — MySQL/PostgreSQL for structured storage; (5) Data Preprocessing Module — cleaning, normalization, and feature extraction; (6) Machine Learning Engine — model training, evaluation, and prediction; (7) Reporting and Visualization Module — charts, dashboards, and downloadable reports.

4.2 Data Model Design

The data model consists of five entities: Admin (admin_id, name, email, password), Vendor (vendor_id, stall_name, contact), Product (product_id, name, price, vendor_id), Sales (sale_id, product_id, quantity, total_price, date_time), and Prediction (prediction_id, product_id, predicted_sales, date). The Admin manages vendors; each vendor owns multiple products; sales are recorded per product; predictions are generated per product.
The system workflow is: user logs in → sales data is entered or uploaded → data is stored in the database → preprocessing and cleaning are performed → historical data is analyzed → ML models are trained → future sales predictions are generated → dashboards and reports are displayed to users.

4.3 User Interface Design

The UI is organized into multiple analytical dashboards: (1) Food Court Analysis Page — main interface for importing and viewing the sales dataset with file upload, dataset preview table, and summary panels showing total records, revenue, and stall count; (2) Sales by Stall Dashboard — bar chart showing revenue per stall with dropdown filters; (3) Revenue by Category Dashboard — pie/donut chart representing category contribution (FastFood, Beverages, Desserts); (4) Hourly Sales Analysis Page — line graph plotting revenue against hourly time slots to identify peak and off-peak periods; (5) Recent Transactions Page — real-time transaction table with search, filter, and sort controls; (6) Sales Prediction / Forecasting Page — ML-based forecast graphs comparing actual vs predicted sales with confidence intervals.

4.4 Technology Stack

Frontend: HTML5, CSS3, JavaScript, Bootstrap, Chart.js / Google Charts. Backend: Python with Flask/Django framework, REST APIs. Database: MySQL/PostgreSQL with SQLAlchemy ORM. Machine Learning: Scikit-learn (Linear Regression, Random Forest, SVM), Pandas, NumPy, Matplotlib, Seaborn; trained model stored as .pkl file. Deployment: Apache/Nginx web server, AWS/Heroku cloud hosting, HTTPS with firewall and authentication module.

V. METHODOLOGY

1. Data Collection: Transaction records are collected from food stalls via POS integration or manual CSV/Excel upload. Each record contains Date, Time, Stall Name, Category, Item, Quantity, Total Price, and Payment Type.

2. Data Preprocessing: Raw data is cleaned using Pandas to handle missing values, remove duplicates, and normalize formats. Temporal features (hour, weekday, month, season) are extracted to improve model accuracy.

3. Exploratory Data Analysis: Sales trends are explored across stalls, categories, time-of-day, and payment types using statistical summaries and visual analytics.

4. Model Training and Evaluation: Multiple forecasting models are trained on historical data: SARIMA and Triple Exponential Smoothing as statistical baselines; Prophet for seasonal trend decomposition; SVM and Stacked SVM as the primary ML models. Performance is evaluated using Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE).

5. Prediction and Deployment: The best-performing model — Stacked SVM — is deployed to forecast future hourly and daily sales per stall and category. The trained model is saved as a .pkl file and loaded for real-time prediction via REST API.

6. Dashboard Visualization: Results are presented through interactive charts: hourly sales line graphs, stall-wise bar charts, category revenue charts, and a recent-transactions table.

VI. RESULTS AND DISCUSSION

The system was tested on transaction data from four food stalls — BurgerKing, JuiceBar, PizzaHub, and SweetTooth — spanning three categories: FastFood, Beverages, and Desserts. The dataset covered multiple dates with attributes including date, time, item, quantity, total price, and payment type (Cash, Card, UPI).

Revenue by Category analysis revealed that FastFood generated the highest revenue (above ₹600,000), followed by Desserts (approximately ₹300,000) and Beverages (approximately ₹250,000). Among individual stalls, all four showed comparable total sales in the range of ₹250,000–₹300,000, indicating a well-balanced distribution of customer footfall across vendors.

The Hourly Sales analysis identified peak demand periods between 12:00–14:00 and 18:00–22:00, with near-zero activity in early morning hours (00:00–08:00). This insight enables targeted staffing and inventory planning during high-demand windows.

The AI Sales Prediction module, powered by Stacked SVM, produced accurate next-hour sales estimates (e.g., ₹483.45 in sample tests). Compared to ARIMA, SARIMA, Triple Exponential Smoothing, and Prophet baselines, the Stacked SVM achieved the lowest RMSE and MAPE values, confirming its superiority for multi-vendor food court environments with nonlinear and multi-factor demand patterns.

ISSN: 2581-9429 Volume 6, Issue 8, March 2026 Impact Factor: 8.2

The Payment Type analysis showed UPI and Cash as the dominant payment modes, providing insights for financial planning and digital payment adoption strategies.

VII. SYSTEM OVERVIEW

7.1 Key Features and Functionalities

- Real-Time Sales Data Collection from vendors via web portal or integrated POS systems.
- Centralized Data Management in a single database covering vendors, products, and transactions.
- Machine Learning-Based Sales Prediction using trained models on historical data.
- Interactive Dashboard and Visualization with bar charts, line graphs, and pie charts.
- Multi-Vendor Support for tracking and comparing individual stall performance.
- Secure User Authentication with role-based access control.
- Automated Reporting with daily, weekly, and monthly reports in PDF/Excel formats.
- Alert and Notification System for predicted high-demand or low-stock situations.

7.2 User Roles and Access Levels

Administrator: highest-level access — manages vendor accounts, product categories, system settings, and complete sales reports. Vendor: limited to own stall data — enters sales transactions, views stall-specific reports and demand predictions. Manager/Supervisor: views overall analytics and performance reports for operational decisions without system configuration privileges. System User (Viewer): read-only access to dashboards and reports.

VIII. SECURITY AND SCALABILITY

8.1 Security Measures

The system implements multiple layers of security: (1) User Authentication with secure login credentials; (2) Role-Based Access Control (RBAC) limiting each role to permitted functionalities; (3) Password Encryption with bcrypt hashing; (4) HTTPS Secure Communication for all data transmission; (5) Firewall Protection blocking unauthorized network traffic; (6) Input Validation and Sanitization preventing SQL injection and XSS attacks; (7) Regular Automated Database Backups; and (8) Activity Logging and Monitoring for suspicious behavior detection.

8.2 Scalability Considerations

The system supports horizontal and vertical scaling through: modular layered architecture allowing independent component upgrades; cloud-based deployment on AWS/Google Cloud with auto-scaling; database indexing, normalization, and sharding for large data volumes; load balancing distributing requests across multiple servers; caching of frequently accessed dashboards; and API-based integration enabling connection to future modules such as mobile apps and IoT devices.

IX. CHALLENGES AND LESSONS LEARNED

9.1 Difficulties Encountered

Key challenges during development included: (1) Data Collection — incomplete, manually entered records in inconsistent formats; (2) Data Cleaning — missing values, duplicate entries, and spelling variations in item names; (3) Model Accuracy — insufficient data and improper feature selection in early stages required multiple iterations of algorithm testing and hyperparameter tuning; (4) System Integration — different POS database structures required a data transformation layer; (5) Performance Optimization — growing database size increased prediction latency requiring query indexing.

9.2 Key Learnings and Best Practices

- Importance of clean, structured data and proper preprocessing before ML training.
- Systematic selection of forecasting algorithms based on data characteristics.
- Modular and layered system architecture for maintainability and extensibility.

DOI: 10.48175/IJARSCT-32143

- REST API-based integration for flexibility with external and future systems.
- Continuous testing, logging, and version control throughout the development lifecycle.

X. FUTURE ENHANCEMENTS

The system has significant potential for future enhancement across several dimensions:

- Advanced ML Models: integration of LSTM and Transformer-based time-series forecasting for improved accuracy incorporating weather, holidays, and seasonal events.
- Mobile Application: Android/iOS app for real-time dashboard access and push notifications for vendors and managers.
- IoT Integration: smart billing machines and sensors for automated real-time data capture.
- Customer Behavior Analytics: purchase pattern analysis and AI-based recommendation engine for personalized promotions.
- Microservices Architecture and Big Data tools (Hadoop/Spark) for multi-branch scalability.
- Biometric authentication and blockchain-based transaction recording for enhanced security.
- Multilingual and multi-location support for wider deployment across restaurant chains and malls.

XI. CONCLUSION

This paper presented a comprehensive Food Court Sales Analysis and Prediction System that leverages machine learning to support data-driven decision-making in multi-vendor food court environments. By applying classical statistical models alongside advanced ML techniques — particularly Stacked SVM — the system achieves accurate sales forecasting with lower RMSE and MAPE compared to ARIMA, SARIMA, Triple Exponential Smoothing, and Prophet baselines.

The interactive web dashboard empowers managers, vendors, and administrators with actionable insights on sales trends, demand patterns, revenue by category, and hourly sales behavior. The system reduces manual effort, minimizes food wastage, and optimizes operational efficiency in inventory planning and staffing. With its modular architecture, cloud-ready deployment, and extensible API design, the system provides a solid foundation for future enhancements including deep learning models, mobile access, and IoT integration, making it a valuable and scalable tool for modern food court management.

REFERENCES

[1] A. Krishna, A. V, A. Aich, C. Hegde, “Sales-Forecasting of Retail Stores using Machine Learning Techniques,” IEEE Int. Conf. on Computational Systems and Information Technology for Sustainable Solutions, 2019.

[2] S. Cheriyan, S. Ibrahim, S. Mohanan, S. Treesa, “Intelligent Sales Prediction Using Machine Learning Techniques,” Int. Conf. on Computing, Electronics and Communications Engineering, 2019.

[3] C. Lu, S. Feng, J. Huang, X. Ye, “A Multi-Task Prediction Framework for Sales Prediction,” ICCECE, IEEE Xplore, 2021, pp. 194–198.

[4] I. Chen, C. Lu, “Sales forecasting by combining clustering and machine-learning techniques,” 2019.

[5] X. Dairu, Z. Shilong, “Machine Learning Model for Sales Forecasting by Using XGBoost,” ICCECE 2021, IEEE.

[6] K. Punam, R. Pamula, P. K. Jain, “A Two-Level Statistical Model for Big Mart Sales Prediction,” GUCON, IEEE, 2018, pp. 617–620.

[7] T. Tanizaki, T. Hoshino, T. Shimmura, T. Takenaka, “Demand forecasting in restaurants using machine learning and statistical analysis,” Procedia CIRP, vol. 79, pp. 679–683, 2019.

[8] X. Ma, Y. Tian, C. Luo, Y. Zhang, “Predicting future visitors of restaurants using big data,” ICMLC, IEEE, vol. 1, pp. 269–274, 2019.

[9] W. Huang, Q. Zhang, “A Novel Trigger Model for Sales Prediction with Data Mining Techniques,” Data Science Journal, 2019.

[10] B. Zhang, R. Tan, C. Lin, "Forecasting of e-commerce transaction volume using a hybrid of ELM and improved moth-flame optimization," Applied Intelligence, 2020.
[11] S. N. Gunjal, S. K. Yadav, D. B. Kshirsagar, "A Distributed Item Based Similarity Approach For Collaborative Filtering On Hadoop Framework," Advances in Parallel Computing, IOS Press, Mar. 2020, pp. 407–415.
[12] A. S. Weigend, N. A. Gershenfeld, Time Series Prediction: Forecasting the Future and Understanding the Past, Addison-Wesley, 2019.