Shri Murlidhar Karwa Knowledge Resource Centre (DSIMS & DSGS)

The Management Quest
Online ISSN: 2581- 6632
Vol.7, Issue 2, April – September 2025

Forecasting Daily Potato Prices in a Mumbai Mandi Using Statistical and Machine Learning Techniques

*** Parth Laijawala** **** Maneesh Gupta**

ABSTRACT

Accurate forecasting of agricultural commodity prices is essential for ensuring food security, supporting farmer income, and enabling effective market planning. This research focuses on the prediction of daily wholesale potato prices in a Mumbai-based mandi using a combination of statistical and machine learning (ML) models. The study utilizes historical price data from January 2021 to December 2024, with the models trained on data from 2021 to 2023 and tested on out-of-sample data from 2024.

Four models were employed: the machine learning models XGBoost and LightGBM, and the statistical models SARIMA and Prophet. The ML models incorporated feature engineering techniques such as lag variables, rolling averages, and calendar features, while the statistical models used the raw time-series data. Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Squared Error (RMSE) were used to assess forecast accuracy.

The results indicate that LightGBM outperformed all other models, achieving the lowest MAPE of 11.90%, followed closely by XGBoost. In contrast, SARIMA and Prophet recorded higher error rates, highlighting the limitations of purely statistical approaches in capturing real-world price volatility. This study emphasizes the importance of adopting feature-rich, data-driven models for short-term agricultural price forecasting and provides valuable insights for market stakeholders, policymakers, and supply chain planners.

Keywords: *Potato price forecasting, Time series analysis, SARIMA, LightGBM, XGBoost, Prophet model.*

INTRODUCTION

Agricultural price forecasting plays a crucial role in ensuring food security, minimizing supply chain inefficiencies, and supporting data driven policy formulation. Among key crops, potato holds high economic and nutritional value in India. It is one of the most important food crop in the world and serves as a staple in Indian diets across regions. India is currently the second-largest producer of potatoes globally, after China, contributing over 50 million Metric Tonnes annually[1]. The major potato-producing

**** Student, Durgadevi Saraf Institute of Management Studies, Mumbai***
***** Assistant Professor, Durgadevi Saraf Institute of Management Studies, Mumbai***

states include Uttar Pradesh, West Bengal, Bihar, Gujarat, and Madhya Pradesh. However, prices in urban Mandis like Mumbai are highly sensitive to supply-demand shifts, transportation delays, and local market disruptions. This study focuses on forecasting daily wholesale prices of potatoes in a Mumbai-based agricultural market (Mandi), primarily the Vashi APMC market (mandi), where price fluctuations significantly impact both producers and consumers.

Potato is a staple crop in India, and its price volatility directly affects both consumers and producers. Timely and accurate price forecasting can help farmers make informed harvesting and selling decisions, reduce post-harvest losses, and aid government bodies in market regulation. Traditional forecasting methods often struggle with high-frequency, non-linear data. Therefore, this study explores the use of both statistical and machine learning models to improve forecasting accuracy.

The research uses daily potato price data from January 1, 2021, to December 31, 2024, encompassing multiple cycles of seasonality, festivals, climatic variations, and post-pandemic recovery effects. We did not use data prior to 2021 because there were significant gaps in the 2020 dataset, largely due to disruptions caused by the COVID-19 pandemic and nationwide lockdown, which affected both data availability and market consistency. The dataset is divided into training data (2021–2023) and testing data (2024) to evaluate how well each model generalizes to unseen future prices.

This study is centered on APMC market a wholesale Mandi in Mumbai, which is a metropolitan hub known for high volume consumption. Mumbai's market dynamics are shaped by both urban demand and interstate supply chains, especially from major producing states like Uttar Pradesh and Gujarat. Price trends in Mumbai also influence regional retail prices and often act as a reference point for nearby markets, therefore emphasizing of a making a forecasting exercise.

The research follows a comparative modeling framework, applying both statistical models (SARIMA, Prophet) and machine learning models (XGBoost, LightGBM) to the same dataset. Each model is evaluated using widely accepted error metrics like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) to objectively assess forecasting accuracy. The comparative analysis highlights the strengths and weaknesses of each modeling approach in capturing the nuances of high-frequency agricultural price data.

LITERATURE REVIEW

Forecasting agricultural commodity prices, particularly for perishable crops like potatoes, has long been a subject of research interest due to its implications for food security, farmer income stability, and market efficiency. Several studies have explored a wide range of statistical, machine learning (ML), and deep learning (DL) methods to model and predict these volatile price patterns.

Badal et al. (2022) applied the ARIMA model to forecast potato prices in major Indian states like Uttar Pradesh, West Bengal, and Bihar. Their study confirmed that ARIMA could produce reliable short-term forecasts when applied to weekly data. However, the authors noted the limitations of ARIMA in capturing complex non-linear patterns inherent in agricultural price movements. Similarly, earlier works have acknowledged the simplicity and interpretability of time-series models like ARIMA and SARIMA but also recognized their inability to adjust to structural breaks, sudden shocks, or high volatility without additional intervention.

Recent advancements in computational power have led to the adoption of more flexible and adaptive models. Nayak et al. (2024) evaluated various deep learning architectures, including LSTM, GRU, CNN, and the relatively novel N-BEATS model, for forecasting weekly potato prices in the Farrukhabad market of Uttar Pradesh. Their findings highlighted that N-BEATS consistently outperformed traditional neural networks and time series models across all accuracy metrics (MAE, RMSE, and MAPE), suggesting a stronger capacity to capture complex temporary patterns in agricultural data.

In a broader context, Zhao et al. (2024) proposed a hybrid framework called VPF-MoE (Vegetable Price Forecasting using Mixture of Experts), which combines large language models (LLMs) with deep learning components. Their ensemble approach dynamically selected the best prediction model based on the characteristics of the vegetable price series. Though their work primarily focused on vegetables like cauliflower and eggplant in the Chinese market, it underscored the growing interest in model ensembling for price prediction, especially when facing heterogeneity across timeframes and commodities.

Qiao et al. (2024) used an ARMA-GARCH framework to capture both the trend and volatility in green onion prices in Korea. They by decomposing price data into trend, seasonal, cyclical, and irregular components using advanced techniques like the Christiano–Fitzgerald (CF) filter and CensusX-13 adjustment, were able to isolate and forecast high-volatility periods more effectively. This approach is also relevant to perishable vegetables like potatoes, which often experience price shocks due to weather changes, supply chain disruptions, or festival driven demand.

Kumar et al. (2022) applied the ARIMA model to analyze and forecast monthly wholesale potato prices in Agra over a nine-year period. Their analysis revealed clear seasonal trends, with prices remaining low from November to January and prices start increasing up from February and peak in November. The ARIMA (2,1,1) model showed strong forecasting ability, supported by low MAPE and MSE values. However, while effective for short-term prediction, the model's performance was sensitive to the data's seasonal and trend components, which may limit its adaptability during periods of sudden market shocks or irregular fluctuations.

Mithiya et al. (2019) utilized the SARIMA model to forecast potato prices in Hooghly district, West Bengal, based on a decade of monthly price data. The selected SARIMA (1,1,0) (4,1,0) model captured both the seasonal and trend components effectively, with prices generally low between January and April and highest in November. Diagnostic checks confirmed the model’s robustness, with favorable error metrics such as low RMSE and MAPE. The study emphasized the model’s utility for planning sale timing but also acknowledged that time-series models like SARIMA, though accurate in stable environments, may require enhancements when market dynamics shift rapidly.

Kumar et al. (2019) developed a pre-season crop price forecasting system that uses historical price data, weather parameters, and sowing dates to predict potato and other crop prices both annually and monthly. By applying ARIMA on rolling means, standard deviation checks, Dickey-Fuller tests, and autocorrelation analyses, they demonstrated up to 95 % confidence in their pre-sowing potato forecasts. Their work highlights the promise of integrating agronomic factors into time-series models but notes that predictive accuracy hinges on the availability and granularity of daily price and climate data, suggesting room for improvement as more frequent observations become available.

Pavithra et al. (2024) evaluated four exponential-smoothing variants (SES, Holt’s, Holt–Winters) alongside ARIMA on potato prices data of the period 2010–2023 in Karnataka’s Kolar market. The Triple Exponential Smoothing (Holt–Winters) model delivered the best fit with results, MAPE of 0.12 % and RMSE of 207, therefore, easily outperforming ARIMA on both training and test splits. Their forecasts for 2024 showed remarkably stable monthly prices, suggesting that smoothing methods excel when seasonality and trend are well-behaved, but they caution these methods may falter if structural breaks or sudden shocks occur outside the historical window.

Jha et al. (2013) compared pure ARIMA, time-delay neural networks (TDNN), and a hybrid ARIMA–TDNN approach on monthly soybean and rapeseed mustard prices. They showed that neural networks captured nonlinear patterns and turning points better than linear models, and that the ARIMA–TDNN hybrid outperformed both individual methods in series exhibiting strong nonlinear dynamics. Evaluation throgh

RMSE, MAD, and correct-direction metrics confirmed the hybrid model's superiority for complex series, though purely linear series saw little benefit, underscoring the importance of pre-testing for nonlinearity when choosing forecasting techniques.

Kumar et al. (2024) applied ARIMA models to monthly onion prices in three Gujarat wholesale markets, Mahuva, Ahmedabad, and Gondal over the period 2004 to 2020. The best-fit models for each market were determined to be ARIMA (3,1,2), ARIMA (2,1,1), and ARIMA (2,1,2), respectively. The post-sample projections for early 2021 showed MAPE values of 21.77 %, 22.99 %, and 27.51 %. Their results underscore how onion price dynamics vary significantly even within one state, highlighting the importance of market-specific model tuning. However, by relying solely on univariate ARIMA, the approach may miss exogenous shocks (e.g., weather disruptions, policy changes), suggesting that incorporating external predictors could further improve forecast robustness.

The reviewed studies collectively demonstrate the evolution of forecasting models from traditional univariate statistical tools to complex, adaptive AI powered systems. However, most of the research has either focused on weekly or monthly prices, with limited exploration at the daily level, particularly for potato prices. This gap restricts the ability of stakeholders to make short-term, data-driven decisions. Our study addresses this gap by focusing on daily price forecasting of potatoes, which enhances the timeliness and granularity of predictions. Furthermore, unlike prior studies that typically rely on a single modeling approach, we compare both statistical (Prophet and SARIMA) and machine learning techniques (XgBoost and LightGBM), providing a broader perspective on model performance and predictive accuracy in the context of agricultural commodity pricing.

OBJECTIVES

The objectives of our study is twofold:

1. To develop statistical and machine learning models to forecast daily wholesale potato prices in the Vashi APMC market.
2. To evaluate and compare the performance of the statistical and machine learning forecasting models.

DATA AND METHODOLOGY

1) Data Description

This study utilizes a daily time-series dataset comprising wholesale (model) prices of potatoes from a major agricultural market (Mandi) located in Mumbai. The time span of the data ranges from 1st January 2021 to 31st December 2024, providing a robust sample that captures, festival-driven demand spikes, seasonal trends, and macroeconomic shocks such as inflation and climate disruptions. The primary source of the data was the **National Horticulture Board (NHB)**, a government agency under the Ministry of Agriculture & Farmers Welfare that is renowned for keeping accurate and up-to-date market price records of horticultural products in all Indian states.

The dataset was structured in a simple two column format as follows:

- **Date**: Representing the calendar date of the observation.
- **Price:** Representing the daily wholesale (model) price of potato (in Indian Rupees per Quintal)

This structure remained consistent across all forecasting models used in the study, except for the Prophet model, which required column names to be modified to ds (for date) and y (for price) as per the model's input requirements. This changes however, did not affect the underlying data content or structure.

The final dataset contained 1,461 observations, corresponding to the number of days over the four-year period. However, during initial exploration of data, it was observed that several dates had missing price values, mainly due to non-operational market days such as Sundays, public holidays, or administrative reporting delays. So, as we know, missing data is a critical issue in time-series analysis, especially for models like SARIMA and Prophet, which assume continuous and evenly spaced intervals. To address this issue, data imputation techniques were used. This approach ensured the continuity and integrity of the dataset without introducing artificial volatility or abrupt structural changes.

Table 1: Summary Statistics of Daily Potato Prices by Year (₹/quintal)

Year	Min Price	Max Price	Median Price
2021	1000	2050	1250
2022	950	2150	1650
2023	900	2200	1150
2024	1000	2750	2250

This table indicates that potato prices in the Mumbai mandi have experienced significant intra- and inter-annual fluctuations. Notably, the median price in 2024 (₹2250/quintal) shows a substantial increase compared to prior years, suggesting a sharp upward trend or supply-side shock. The price floor remained relatively stable (₹900–₹1000), while the price ceiling increased from ₹2050 to ₹2750, reflecting heightened volatility and supporting the need for robust forecasting mechanisms.

These insights guided the selection of forecasting models capable of handling non-linearity, trend shifts, and seasonality, such as SARIMA, Prophet, XGBoost, and LightGBM, each evaluated under the same data conditions for consistent performance comparison.

2) Feature Engineering

Feature engineering plays a critical role in enhancing the predictive performance of machine learning models, particularly in time-series forecasting where capturing temporal dependencies, seasonality, and structural patterns is essential.

For this study, no additional features were engineered for the statistical models, namely SARIMA and Prophet. These models were trained directly on the original univariate time-series data, i.e., the daily wholesale (model) price of potatoes. The model's internal mechanisms inherently account for trend and seasonality, making feature expansion unnecessary and, in some cases, counterproductive.

In contrast, the machine learning models like LightGBM and XGBoost, required additional input features to model temporal dependencies and capture hidden patterns more effectively. The following categories of features were engineered and incorporated into the training dataset:

Lag Features: Lagged versions of the price variable were created to capture short-term dependencies. These included lag_1 and lag_7, representing the price values of the previous 1 to 7 days, respectively. These features help the models recognize autoregressive behavior in price fluctuations.

Rolling Window Statistics: To incorporate medium-term trend information, rolling mean features were added. These included 3-day, 7-day, and 14-day moving averages of past prices. Such features smooth out short-term volatility and help the models understand underlying price trends.

Calendar-Based Features: Temporal calendar attributes were extracted from the Date column to help the models detect recurring patterns and seasonality. The following features were used:

- Day of the Week (day_of_week): 0 = Monday to 6 = Sunday
- Month of the Year (month): 1 = January to 12 = December
- Is the day a weekday or weekend (is_weekend): a binary function – 1 if Saturday or Sunday, 0 otherwise

These engineered features allowed the machine learning models to learn both short-term fluctuations and seasonal cycles more effectively, which would not be possible using the raw price series alone. No scaling or normalization was required, as both

LightGBM and XGBoost are tree-based algorithms that are insensitive to the feature scale.

The same set of features was used for both machine learning models to ensure a fair comparison in the performance evaluation phase.

3) Model Description

This study employs a combination of statistical and machine learning models to forecast daily wholesale potato prices in a Mumbai-based agricultural market. The objective is to compare the predictive capabilities of traditional time-series methods with modern, data-driven approaches, particularly in handling daily-level fluctuations. Each model utilized is described in depth below:

3.1) SARIMA (Seasonal AutoRegressive Integrated Moving Average)

The SARIMA model is a traditional statistical method for time-series forecasting that builds on the ARIMA model by including seasonality. It is particularly well-suited for datasets where patterns repeat over fixed intervals, such as monthly or yearly cycles. SARIMA works by combining three core elements:

- **Autoregression (AR):** Models the relationship between an observation and a certain number of lagged observations.
- **Differencing (I):** is a method for making a time series stationary by removing trends.
- **Moving Average (MA):** The relationship between an observation and a residual error resulting from a moving average model applied to lagged data.

To account for seasonality, SARIMA adds seasonal terms to each of these components. This enables the model to capture repeating patterns at fixed time intervals, such as weekly or monthly price changes in agricultural commodities.

SARIMA is widely used in economic and agricultural forecasting due to its interpretability and strong performance in datasets with clear seasonal trends. In this study, SARIMA was applied directly to the raw price series. The model learns from past trends and seasonal cycles to predict future prices, particularly well-suited to understanding structured and recurring patterns in the potato price data.

3.2) Prophet

Prophet, developed by Facebook’s Core Data Science team, is a robust and flexible time-series forecasting model designed to handle data with strong seasonal effects and historical trend changes. Unlike traditional models, Prophet allows for an intuitive decomposition of the time series into three components:

- **Trend:** Measures the long-term increase or decline in a time series.
- **Seasonality:** refers to periodic fluctuations that occur at regular intervals, such as weekly or annual cycles.
- **Holidays or Events:** Allows for the inclusion of known events that may impact the time series, such as festivals or policy changes.

One of Prophet’s key advantages is its ability to automatically detect changepoints, moments in time where the trend shifts significantly, and adjust the forecast accordingly. It employs an additive model in which non-linear trends are combined with seasonal and holiday impacts.

Prophet is particularly useful in business and policy-related forecasting tasks due to its ease of use, minimal parameter tuning, and strong performance on irregular or noisy data. In this study, Prophet was trained on the daily price series of potatoes to identify key turning points and capture both weekly seasonality and long-term trend shifts.

3.3) LightGBM (Light Gradient Boosting Machine)

LightGBM is a gradient boosting framework developed by Microsoft that is designed for fast, scalable, and high-accuracy performance. It is based on decision tree algorithms and operates by sequentially developing models, with each new model focusing on fixing prior faults. Unlike standard boosting approaches, which grow trees level-wise, LightGBM grows trees' leaf-wise, which frequently results in higher accuracy.

The model is especially efficient with large datasets and supports features like categorical variable handling, missing value handling, and parallel training. Its **advantages** include:

- Faster training speed and lower memory usage
- Better accuracy
- Ability to handle large-scale data

In the context of this study, LightGBM was used to model the relationship between the daily potato prices and several engineered features such as lag variables, rolling means, and calendar indicators. The model's ability to learn complex, non-linear relationships from structured data made it well-suited for this forecasting task. Furthermore, LightGBM’s feature importance tools provided insights into which factors had the most predictive power.

3.4) XGBoost (Extreme Gradient Boosting)

XGBoost is another popular machine learning algorithm that uses gradient boosting decision trees. Known for its predictive power, speed, and regularization techniques, XGBoost has become a benchmark model in structured data forecasting competitions and academic research.

Like LightGBM, XGBoost builds models in a sequential manner, with each new tree attempting to minimize the residual errors of the previous ensemble. The key **distinctions** of XGBoost include:

- Regularization (L1 and L2) to avoid overfitting
- Pruning techniques to reduce complexity
- Advanced treatment of missing values and sparse data.

In this study, XGBoost was applied using the same engineered features as LightGBM. The model effectively captured both short-term lags and seasonal influences, allowing it to adapt to complex pricing behavior in the agricultural domain. Despite being computationally intensive, XGBoost offered strong performance in terms of forecast accuracy and robustness.

Table 2: Summary of Forecasting Models Used in the Study

Model	Type	Key Strengths	Input Requirements	Feature Engineering Needed
SARIMA	Statistical Time Series	Captures seasonality and trend; interpretable	Univariate time series	No
Prophet	Additive Time Series	Handles changepoints, holidays; easy to tune and interpret	Date (ds) and value (y)	No
LightGBM	Machine Learning	Fast, scalable, handles non-linear patterns	Tabular data with time features	Yes
XGBoost	Machine Learning	High accuracy, robust to overfitting, strong handling of noise	Tabular data with time features	Yes

Each of the four models was evaluated using a common training and testing structure to ensure fairness and comparability in their predictive outcomes. The next section details the train-test split strategy adopted for this purpose.

4) Train-Test Split Strategy

To evaluate the predictive performance of the forecasting models on unseen data, a chronological train-test split was adopted. This approach is essential in time-series forecasting, where data exhibits strong temporal dependencies and traditional random shuffling methods are not applicable.

The full dataset spans from 1st January 2021 to 31st December 2024, comprising four full calendar years of daily wholesale potato prices. The dataset was split as follows:

- Training Period: 1st January 2021 to 31st December 2023
- Testing Period: 1st January 2024 to 31st December 2024

This split ensures that the models learn from past historical data (three complete years) and are evaluated on future data (a full unseen year), thus simulating a real-world forecasting scenario. The division also aligns with best practices in time-series modeling, which emphasize the importance of maintaining temporal order during validation.

The decision to use 2024 as the test year was based on the need to:

- Evaluate model robustness over a full seasonal cycle
- Capture forecasting performance across different market conditions (festivals, harvest cycles, lean periods)
- Maintain consistency across all models for fair comparison

All four models—SARIMA, Prophet, LightGBM, and XGBoost, were trained exclusively on the 2021–2023 dataset and then used to generate out-of-sample forecasts for the entire year of 2024. This static train-test split was deemed sufficient given the dataset's size and the study's focus on model comparison, not real-time retraining.

The evaluation of the forecasts, discussed in the next section, was performed by comparing the predicted prices for 2024 against the actual observed prices using standard error metrics.

5) Evaluation Metrics

To objectively assess the performance of the forecasting models, the study employs three widely used error metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE). These metrics provide a comprehensive evaluation of model accuracy, penalizing different types of errors and helping to understand both the average and relative prediction performance.

Each metric was computed by comparing the actual daily wholesale prices of potatoes in 2024 (the test set) with the corresponding model-generated forecasts from SARIMA, Prophet, LightGBM, and XGBoost. The formulas and rationale for each metric are outlined below.

1. Mean Absolute Error (MAE)

$$MAE = \frac{1}{n}\sum_{t=1}^{n} | \, y_t - \widehat{y_t} \, |$$

MAE measures the average magnitude of the absolute errors between the actual and predicted values. It is easy to interpret and is **less sensitive to large outliers** compared to RMSE. A lower MAE indicates better model performance.

2. Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} (\, y_t - \widehat{y_t} \,)^2}$$

RMSE penalizes larger errors more heavily due to the squaring of residuals. It is particularly useful when large deviations are undesirable. RMSE is expressed in the same units as the target variable (₹/quintal), which makes it intuitive to interpret in the context of price prediction.

3. Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{100}{n}\sum_{t=1}^{n} | \, \frac{y_t - \widehat{y_t}}{y_t} \, |$$

MAPE expresses forecast error as a **percentage of the actual values**, making it scale-independent. It is especially helpful for comparing performance across different time periods or datasets. However, MAPE can be distorted when actual values are very small.

All models were evaluated using the same test set (daily prices for the year 2024) to ensure consistency and comparability. The metrics were calculated using built-in functions from **Python libraries** such as **scikit-learn and statsmodels**, depending on the model type.

RESULTS

This section presents the results of the four forecasting models used, SARIMA, Prophet, LightGBM, and XGBoost which were evaluated on the daily wholesale potato prices for the test period of 1st January 2024 to 31st December 2024. The models were assessed using three standard evaluation metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE), as described in above section.

Table 3: Forecasting Performance of Models on 2024 Daily Potato Prices

Model	MAE (₹/quintal)	RMSE (₹/quintal)	MAPE (%)
LightGBM	279.28	362.88	11.90%
XGBoost	281.22	364.06	12.02%
SARIMA	795.69	919.02	34.92%
Prophet	852.87	975.90	37.50%

1) Visual Interpretation of Forecast Accuracy

To supplement the numerical results, actual vs predicted line plots were generated for each model across the test period (2024). These visualizations help assess how closely each model's forecast tracks the real price movements throughout the year, and reveal strengths and weaknesses not always captured by numerical error metrics alone.

The forecast generated by the Prophet model in figure 1 below shows that the model remains relatively flat and smoothed, failing to capture the real-world fluctuations seen in the actual price trend. While Prophet is capable of modeling seasonal and trend components, its limited responsiveness to sharp price changes results in significant underfitting. This is particularly evident during the mid-year and end-of-year peaks, where the model underestimates the true values. The widening confidence interval also reflects increasing uncertainty over time, typical of additive models that assume smooth future evolution.

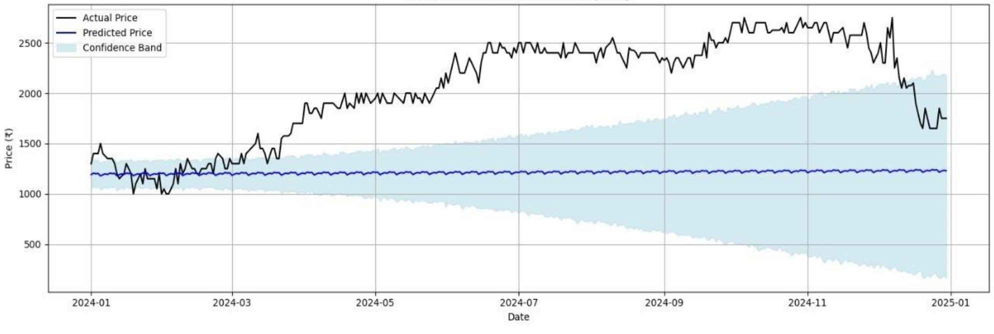

Figure 1: Prophet – Actual vs Predicted with Confidence Interval

The graph in Figure 2 below presents a comparison between the actual daily potato prices observed in the Mumbai Mandi for the year 2024 and the forecasted prices generated using the SARIMA model, along with the 80% and 95% prediction intervals.

The black line represents the actual price trend, which exhibits strong seasonality and volatility throughout the year. There is a noticeable price increase during the middle of the year, followed by a sharp decline toward the end.

In contrast, the red line illustrates the forecasted prices from the SARIMA model, which remain almost flat across the entire forecast horizon. This flatness indicates that the SARIMA model failed to capture the dynamic price movements, seasonal fluctuations, and non-linear trends observed in the actual data.

The shaded regions denote the confidence intervals:

- The dark blue band represents the 80% confidence interval.
- The light blue band represents the 95% confidence interval.

Although the forecast includes these uncertainty bands, the actual prices consistently fall outside the predicted range, especially during the periods of sharp increase and decline. This signifies poor forecast accuracy and highlights the model's limited ability to adapt to sudden market shifts.

Moreover, the widening of the confidence intervals over time suggests that the SARIMA model becomes increasingly uncertain in its long-term predictions. This is a known limitation of univariate time series models like SARIMA, which rely solely on historical price values and do not incorporate external or engineered features such as calendar effects, weather conditions, or lagged trends.

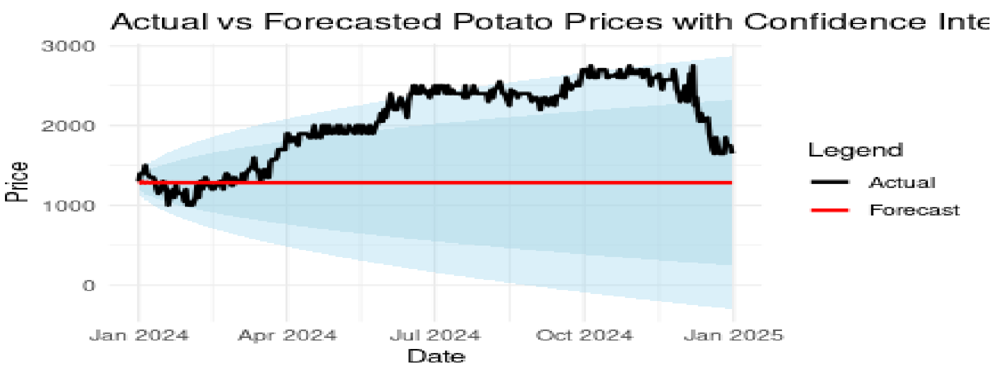

Figure 2: SARIMA – Actual vs Predicted

The graph in figure 3 below reflects that **Extreme Gradient Boosting (XGBoost)** also exhibits strong predictive power, closely mirroring the overall pattern of actual potato prices throughout 2024. Like LightGBM, XGBoost benefits from the inclusion of timeaware features such as previous day prices, rolling windows, and temporal indicators (e.g., day of the week, month). These features enable it to capture both autoregressive dependencies and seasonal effects.

XGBoost demonstrates particularly good performance during steady market phases, such as the mid-year period (June to August), where it maintains a tight alignment with observed values. In these periods, the model's low variance and strong regularization allow it to make precise, stable predictions with minimal noise.

However, in periods of heightened volatility, such as late Q3 and early Q4, the model exhibits a slight tendency to underpredict during extreme spikes. This behavior is not uncommon for tree-based ensemble methods that prioritize average trend capture over outlier sensitivity. Nonetheless, XGBoost still correctly anticipates the direction and general shape of the trend, which is critical for operational forecasting in agricultural markets.

Compared to LightGBM, XGBoost appears to be marginally more conservative, possibly due to its different boosting mechanics (level-wise vs. leaf-wise tree growth). While this makes it slightly less reactive to abrupt shifts, it also reduces the risk of overfitting, leading to robust performance across most timeframes.

With a MAPE of 12.02%, XGBoost stands as a close second in overall model performance. Its consistent accuracy and reliable generalization make it a highly effective forecasting tool, particularly in scenarios where moderate price fluctuations are expected, and stability is preferred over high sensitivity.

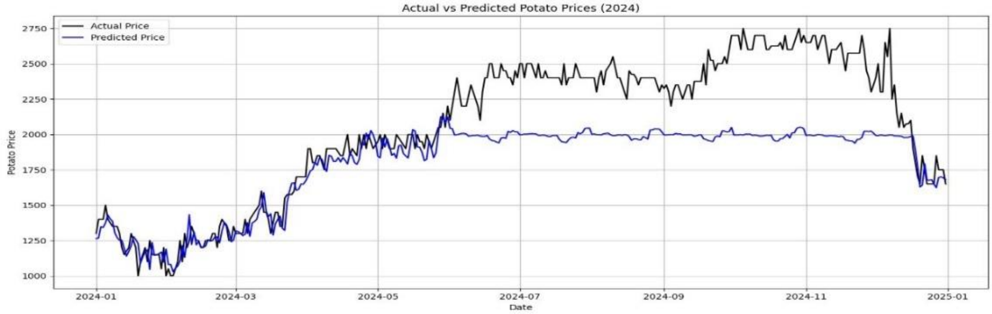

Figure 3: XGBoost – Actual vs Predicted

The graph in Figure 4 below shows that the Light Gradient Boosting Machine (LightGBM) model demonstrated remarkable forecasting accuracy, as evidenced by the close alignment between predicted and actual potato prices across the 2024 test period. Unlike traditional statistical models, LightGBM leverages engineered features such as lagged values, rolling averages, and calendar-based attributes, allowing it to model both short-term fluctuations and long-term seasonal patterns in price behavior.

In the first half of the year, LightGBM responds well to the subtle increases and decreases in prices, capturing the general upward trend around March–April, which may correspond to pre-monsoon demand and changing supply conditions. The model also performs well during the June–July plateau, accurately reflecting periods of market stability.

More notably, in the second half of the year, where the price series becomes more volatile, particularly around September through November, LightGBM still manages to follow the actual trend with high fidelity. Although it slightly underpredicts during sharp price spikes, the magnitude and direction of changes remain largely accurate, indicating that the model effectively learns from historical lag patterns and seasonal cycles.

The model's predictions remain smooth yet reactive, suggesting a good balance between generalization and responsiveness. Unlike SARIMA and Prophet, which often over-smooth or lag behind sudden changes, LightGBM is better at adapting to non-linear transitions due to its boosting framework and ability to handle complex feature interactions.

This performance highlights LightGBM's strength in capturing agricultural price movements, especially in real-world contexts where data is noisy, seasonal, and influenced by multiple indirect factors. Its superior MAPE (11.90%), the lowest among all models tested, further reinforces its suitability for short-term, high-frequency market forecasting in perishable commodity sectors.

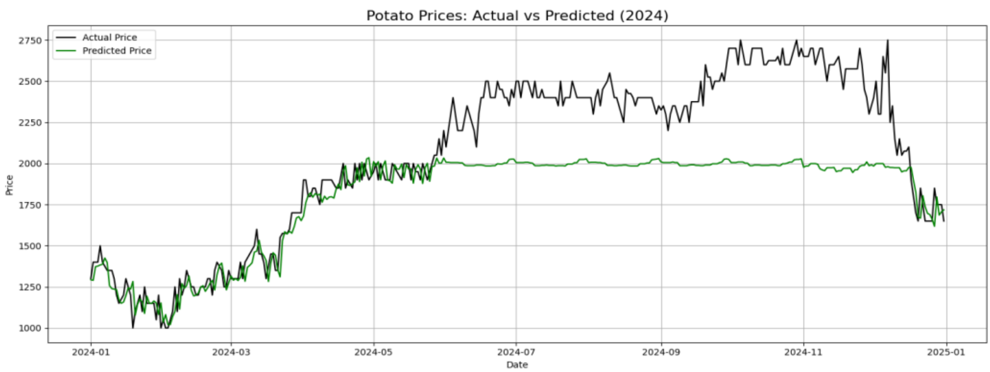

Figure 4: LightGBM – Actual vs Predicted

The results indicate that machine learning models (LightGBM and XGBoost) significantly outperformed the statistical models (Prophet and SARIMA) in all three-evaluation metrics. LightGBM achieved the lowest error rates, with a MAPE of 11.90%, closely followed by XGBoost at 12.02%. This suggests that ensemble tree-based algorithms are better equipped to handle complex patterns in high-frequency, non-linear agricultural price data when supplemented with engineered features.

In contrast, both SARIMA and Prophet, although widely used in univariate time-series forecasting, exhibited higher error values, with MAPE values exceeding 34%. These results show that while statistical models can capture seasonal structures, they may underperform in highly volatile, real-time forecasting scenarios without external features.

RECOMMENDATIONS AND CONCLUSIONS

The results of this study clearly indicate that machine learning models that too specifically LightGBM and XGBoost, outperform traditional statistical methods (SARIMA and Prophet) when it comes to forecasting daily wholesale potato prices in an urban Indian mandi. These findings have practical implications for farmers, policymakers, traders, and supply chain stakeholders involved in the Indian agri-market ecosystem.

LightGBM, in particular, achieved the lowest forecasting error, suggesting that feature-driven models are more effective in capturing non-linear trends, seasonal fluctuations, and abrupt price changes in volatile, high-frequency agricultural markets. XGBoost followed closely, offering robust and consistent performance with slightly conservative predictions. In contrast, SARIMA and Prophet, while effective at modeling seasonal structure, lacked the flexibility to adapt to unexpected price surges or regime shifts, resulting in higher forecast errors.

These insights support a paradigm shift toward hybrid or machine learning–based forecasting frameworks, especially in regions like India, where price fluctuations are driven by a mix of climate patterns, transportation disruptions, festivals, and market interventions. The integration of lag-based, rolling, and calendar features provides a more nuanced understanding of price behavior than univariate historical data alone.

The adoption of ML-based models can enable better decision-making for procurement planning, inventory management, and government price stabilization policies. In the context of Mumbai and similar urban markets, such predictive tools can also inform consumer advisories, reduce wastage, and contribute to more efficient food distribution systems.

1) LIMITATIONS

Despite the encouraging results achieved in this study, several limitations must be acknowledged, which could influence the interpretability and generalizability of the findings:

Limited Feature Scope: The machine learning models were trained using time-derived features such as lag variables, rolling averages, and calendar-based indicators. However, external drivers like rainfall, fuel prices, festival dates, transport disruptions, and other macroeconomic variables, are often influential in agricultural pricing which were not included due to data unavailability. Their inclusion could potentially enhance forecast precision, especially during high-volatility periods.

Missing Data Periods: One of the critical constraints faced during dataset preparation was missing historical data. Notably, there was a significant data gap from 22nd March to 26th May 2020, a period coinciding with India's first COVID-19 lockdown, where markets were either closed or reporting was inconsistent. Additionally, in the year 2017, data was missing for nearly four consecutive months starting from June. Due to these discontinuities, the dataset used in this study was limited to January 2021 to December 2024, thereby restricting the training window and limiting the model's ability to learn from long-term patterns, particularly during the months of March and April, which are typically transitional periods in the crop supply cycle.

Single-Market Limitation: The analysis is based solely on data from one wholesale market in Mumbai, which, although representative of urban consumption centers, may not fully capture price dynamics across rural or semi-urban mandis. Hence, caution is advised in generalizing these findings to different regions without further validation.

2) FUTURE SCOPE OF RESEARCH

This study opens several avenues for further exploration. Future research can benefit from the inclusion of external factors such as weather patterns, festival calendars, fuel prices, and policy interventions to improve model accuracy, especially during volatile periods. Expanding the dataset to include multiple markets across different regions could enhance the generalizability of results and support broader policy applications. Additionally, exploring hybrid models that combine statistical techniques with deep learning architectures (e.g., LSTM, N-BEATS) may offer improved adaptability to complex, real-world pricing behavior.

REFERENCES

- Badal, P. S., Kamalvanshi, V., Goyal, A., Kumar, P., & Mondal, B. (2022). Forecasting potato prices: application of ARIMA model. *Economic Affairs*, (4), 491-496. https://doi.org/10.46852/0424-2513.4.2022.14
- Jha, G. K., & Sinha, K. (2013). Agricultural price forecasting using neural network model: An innovative information delivery system. *Agricultural Economics Research Review*, *26*(2), 229-239. https://doi.org/10.22004/ag.econ.162150
- Kumar, M., Sharma, R. K., Bhopala, U. D., & Suthar, B. (2024). Forecasting prices of onion in major wholesale markets of Gujarat. International Journal of Statistics and Applied Mathematics, 9(1S), 22-29.
- Kumar, S., Chaddha, H., Kumar, H., & Singh, R. K. (2019). Price forecasting system for crops at the time of sowing. *Int Res J Eng Technol (IRJET)*, *6*(4).
- Kumar, T. M., Munirajappa, R., & Surendra, H. S. (2011). Application of seasonal ARIMA model for forecasting monthly prices of potato in Bangalore market. *Editorial Committee*, *45*(4), 778-782.
- Mithiya, D., Mandal, K., & Datta, L. (2019). Forecasting of potato prices of Hooghly in West Bengal: time series analysis using SARIMA model. *Int J Agric Econ*, *4*(3), 101. https://doi.org/10.11648/j.ijae.20190403.13
- Nayak, G. H., Alam, M. W., Avinash, G., Singh, K. N., Ray, M., & Kumar, R. R. (2024). N-BEATS deep learning architecture for agricultural commodity price forecasting. *Potato Research*, 1-21. https://doi.org/10.1007/s11540-024-09789-y
- Pavithra, V. (2024). A statistical approach to forecasting potato price in Kolar market of Karnataka. *Int J Statistic Appl Math*, *9*(2), 32-36.
- Qiao, Y., & Ahn, B. I. (2024). Volatility analysis and forecasting of vegetable prices using an ARMA-GARCH model: An application of the CF filter and seasonal adjustment method to Korean green onions. Agribusiness. https://doi.org/10.1002/agr.21958
- Zhao, S., Zhang, Y., Xu, Q., Yang, Y., & Zhang, T. (2024). Vegetable price forecasting using a mixture-of-experts framework with large language models. Agriculture, 15(1), 162. https://doi.org/10.3390/agriculture15020162