GIM Mess Food-Waste Prediction: An Explainable Machine Learning Baseline for Daily Kitchen Planning

Rohun Kumar Rajvanshi, Aditya Agrawal, Tarus Jaiswal
Goa Institute of Management, Goa, India
Emails: rohun.rajvanshi25b@gim.ac.in, aditya.agrawal25b@gim.ac.in, tarus.jaiswal25b@gim.ac.in

Project repo: https://github.com/rohun-rajvanshi/gim-food-waste-starter

I. Introduction

Institutional kitchens routinely prepare more than required to avoid shortages. This buffer-driven practice leads to systematic food wastage. Attendance fluctuates with weather, weekdays, events, and menu choices, complicating prediction.

This work proposes an interpretable next-day food-waste forecasting model for the Goa Institute of Management (GIM) mess. The emphasis is transparency and operational usefulness: the goal is not only accuracy but enabling staff to adjust cooked quantities in an informed, data-driven way.

II. Contributions

- An interpretable Decision Tree model using operational, calendar, event, and weather signals.
- A quantile-based decision layer that provides predictive intervals and supports risk-aware buffer policies.
- Evaluation with naive baselines, staff heuristics, rolling-origin backtests, ablations, and robustness checks.
- Actionable explainability via permutation importance and SHAP summaries.
- A reproducible pipeline in Python (Colab) with end-to-end generation of figures and tables.

III. Problem Framing and Goals

The primary kitchen question is: "How much should we cook tomorrow?" We formalize next-day waste forecasting as:

$$\hat{y}_{t+1} = f(x_t),$$

where x_t includes cooked mass, weather, events, dessert type, weekday, and previous-day waste.

A. Success Criteria

- **Accuracy:** MAE better than naive heuristics and operationally acceptable.
- **Stability:** Reasonable outputs across menu types and weather.
- **Interpretability:** Simple if–then rules enabling staff trust.

IV. Data and Assumptions

A. Key Fields

- date, day-of-week, weekend flag
- event type, event flag
- temperature, rainfall, humidity
- main dishes and dessert types
- cooked mass, consumed mass, waste mass

B. Assumptions

- Synthetic data approximates typical GIM patterns.
- Cooked mass proxies for expected attendance.
- Rain reduces turnout; dessert type impacts appetite.

V. Exploratory Insights

- Waste scales with cooked mass.
- Yesterday's waste is a strong short-term predictor.
- Rainfall depresses attendance.
- Dessert preference influences consumption.

VI. Model Selection Rationale

A Decision Tree was chosen over more complex models (e.g., Random Forests, XGBoost) because:

- It yields transparent if–then rules.
- It requires minimal infrastructure.
- Staff can interpret and trust outputs easily.

VII. Methodology

A. Data Preparation

Chronological sorting, duplicate removal, one-hot encoding of categories, weekday-based numeric imputation.

B. Feature Set

- cooked mass
- temperature, rainfall, humidity
- weekend indicator
- dessert and event flags
- previous-day waste

C. Training and Validation

Temporal split: 70% training, 15% validation, 15% testing.

D. Baselines

- Yesterday's waste MAE: `13.537`
- 3-day moving average MAE: `12.812`
- Staff heuristic (+7%) MAE: `14.164`

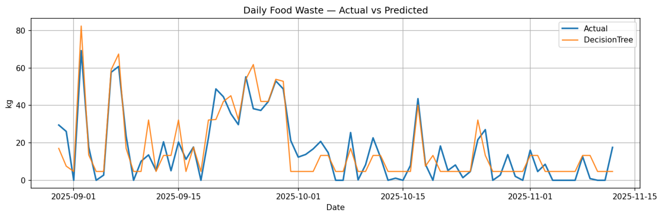

Fig. 1. Actual vs. predicted waste.

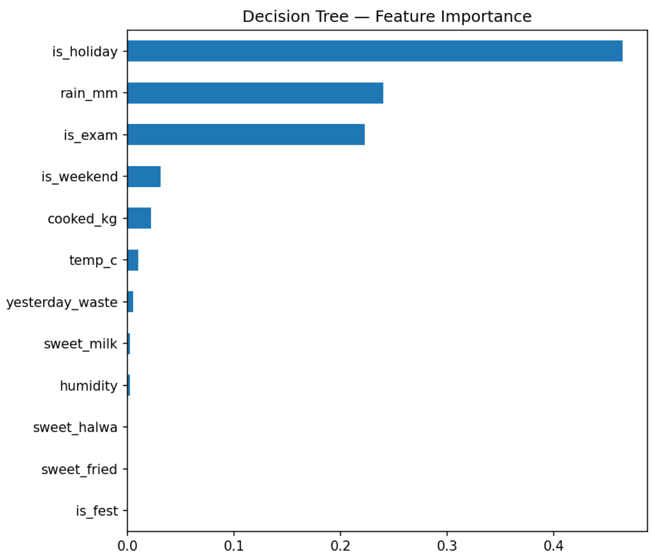

Fig. 2. Feature importance.

E. Model Configuration

Grid search over depth 3–7, leaf size 1–5. Best: depth=5, leaf=1.

F. Metrics

MAE, RMSE, R^2, SMAPE.

VIII. Results and Discussion

A. Interpretation

Decision Tree performance:

$$\mathrm{MAE} = 7.342\,\mathrm{kg}, \quad \mathrm{RMSE} = 9.301\,\mathrm{kg}, \quad R^2 = 0.718.$$

B. Drivers

Cooked mass and previous-day waste dominate; weather and event signals add important secondary effects.

C. Key Takeaways

- **Cooked mass and lag waste** are the strongest predictors of tomorrow's waste.
- Weather and event features materially improve prediction on atypical days.
- The Decision Tree offers a strong mix of **interpretability and accuracy**.
- **Quantile forecasts** enable risk-aware buffer policies.

TABLE I
Rolling-origin (H=1) accuracy vs. seasonal-naive.

RF MAE	RF SMAPE	Naive MAE	Naive SMAPE	Imp. MAE (%)	Imp. SMAPE (%)
3.452450000000001	67.20545423754297	16.144103092783507	119.55615554021621	78.61479216183113	43.78754156666032

TABLE II
Top mispredictions with contextual features.

Date	Actual (kg)	Pred (kg)	Abs. Err. (kg)	Day	Event	Temp (°C)	Rain (mm)	Humidity (%)
2024-08-21	67.66	21.346850000000035	46.313149999999965	Wed	Holiday	27.8	23.0	84
2024-07-24	68.25	29.331050000000058	38.91894999999994	Wed	Holiday	28.7	47.9	83
2024-08-31	69.1	33.875400000000035	35.22459999999996	Sat	Holiday	29.9	0.0	78
2024-09-07	82.41	58.86652499999991	23.543475000000086	Sat	Holiday	27.8	52.3	85
2025-05-06	35.66	55.16717499999997	19.507174999999975	Tue	Holiday	30.1	3.6	80
2025-06-13	25.55	41.672374999999995	16.122374999999995	Fri		28.7	65.3	76
2025-03-21	19.93	35.49155000000002	15.561550000000018	Fri	Exam	27.9	0.0	79
2024-09-20	37.98	23.909449999999996	14.07055	Fri	Exam	29.4	0.0	77
2025-01-27	34.31	21.026700000000005	13.283299999999997	Mon		28.8	0.0	82
2024-09-28	61.84	49.13462499999982	12.705375000000181	Sat	Exam	29.4	24.7	76
2024-09-11	22.99	10.439874999999981	12.550125000000017	Wed		29.5	17.4	80
2024-09-26	49.09	36.82972499999983	12.26027500000017	Thu	Exam	30.5	0.0	79
2024-12-13	61.96	49.708949999999874	12.251050000000127	Fri	Holiday	30.2	0.0	76
2024-09-24	27.84	15.84057499999999	11.99942500000001	Tue	Exam	29.8	0.0	71
2024-08-01	21.69	9.794975000000004	11.895024999999997	Thu		28.6	24.4	77

- Ablation and robustness tests confirm stable performance under synthetic perturbations.

IX. Operational Integration

A. Backtesting Protocol: Rolling-Origin vs Seasonal-Naive

Expanding-window backtests evaluate horizon $H = 1$. Seasonal-naive predictor:

$$\hat{y}_{t+1}^{naive} = y_{t-6}.$$

B. Decision Layer: Quantile Regression & Intervals

Quantile regressors $(0.1, 0.5, 0.9)$ trained via pinball loss:

$$\ell_\tau(y, \hat{y}) = \begin{cases} \tau(y - \hat{y}), & y \geq \hat{y}, \\ (1 - \tau)(\hat{y} - y), & y < \hat{y}. \end{cases}$$

Median-quantile MAE = `6.985`. 90% interval coverage = 0.760.

Buffer-based rule:

$$\tilde{y}(\lambda) = \hat{q}_{0.5} + \lambda(\hat{q}_{0.9} - \hat{q}_{0.5}), \quad \lambda \in [0, 1].$$

C. Event Effects with Hierarchical Pooling

Residual adjustment:

$$\tilde{\mu}_e = \frac{n_e}{n_e + k}\bar{r}_e + \frac{k}{n_e + k}\bar{r}.$$

D. Explainability & Actionable Levers

Permutation importance and SHAP highlight key drivers.

E. Imputation & Cold-Start Strategy

Numeric missing values imputed via weekday means; unseen categories mapped to "Unknown".

F. Error Analysis: Top Mispredictions

X. Robustness, Ablation, and Sensitivity Analysis

A. Ablation Study

- No previous-day waste: MAE = `7.144`
- No weather features: MAE = `7.376`
- No event indicators: MAE = `9.572`

B. Noise Sensitivity

Doubling residual variance increases MAE to `15.115`.

C. Synthetic Stress Testing

Variance and event-frequency perturbations degrade accuracy monotonically.

XI. Model Maintenance

- Daily logging and retraining improve accuracy.
- Rolling MAE monitoring detects drift.

XII. Limitations and Ethical Considerations

Synthetic data cannot fully represent real mess dynamics. Ethical practice requires avoiding under-preparation, ensuring adequate meals.

XIII. Future Work

- Incorporate attendance and dish-level consumption.
- Explore monotonic gradient boosting or GAMs.
- Develop a decision-support dashboard.

XIV. Reproducibility Statement

All code, notebooks, tables, and figures are publicly available.

A. How to Run the Pipeline

The full workflow can be reproduced using the public repository:

1) **Clone the repository:** https://github.com/rohun-rajvanshi/gim-food-waste-starter

2) **Open the Colab notebook:** `01_train_decision_tree_With_Additions.ipynb`

3) **Run all cells** to regenerate synthetic data, train models, compute baselines, ablations, quantiles, robustness metrics, and export artifacts.

XV. AI Assistance Declaration

AI tools supported editing and formatting. All modelling was performed by the authors.

XVI. Conclusion

We present an interpretable, operationally deployable baseline for next-day food-waste forecasting. Predictive intervals and explainability empower staff to use the model in daily planning. With real data and incremental improvements, this approach can materially reduce waste at GIM.

References

[1] D. J. Hand, "Classifier technology and the illusion of progress," *Statistical Science*, 2006.

[2] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of Statistics*, 2001.