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

Rohun Kumar Rajvanshi*, Aditya Agrawal*, Tarus Jaiswal*
*Goa Institute of Management, Sanquelim, Goa, India
Email: 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

***Abstract*—College cafeteria premises typically overcook meals to prevent shortages, which inadvertently results in large amounts of food wastage. This paper is about the Goa Institute of Management (GIM) mess and suggests an explainable machine learning model to predict next-day leftover food. The model combines operational, calendar, and environmental drivers—like planned cooked quantity, weather, weekends, events, dessert type, and wastage from the previous day—to enable kitchen-level planning. The Decision Tree Regressor was chosen for interpretability and ease of deployment, allowing mess staff to visualize decision rules and realize what factors they influence. With a 500-day synthetic dataset that replicates GIM's context, the model performed well (MAE = 6.40 kg, RMSE = 8.10 kg, $R^2 = 0.786$), showing that even a simple method is able to inform sustainable and data-driven meal planning.**

***Index Terms*—Food waste, demand forecasting, decision trees, cafeteria planning, explainable ML, campus operations**

I. Introduction

Mealtimes in institutional kitchens are different every day with weather, calendar date, and weekly rhythms. To ensure availability of food, kitchens tend to over-produce by preparing more than they consume, leading to daily food wastage. This work solves this issue in the GIM mess by forecasting tomorrow's food waste with available data. The goals of the project are to establish a baseline that is simple and reliable for users to work with, reduce data and maintenance effort, and enjoy a workflow conducive to day-to-day usage.

II. Contributions

- Designed an interpretable and transparent Decision Tree Regressor to forecast food waste using few but pertinent features.
- Demonstrated a "what-if" planning model to help decision-makers resize cooking quantities in different situations.
- Published an open and reproducible pipeline on GitHub.

III. Problem Framing and Goals

Business goal: reduce wastage of food without affecting service quality. The machine learning problem is a next-day food wastage in kilograms regression task. Mess manager, kitchen staff, administration, and sustainability office are the primary stakeholders.

The central choice is how much to cook for tomorrow. The model estimates expected waste as a function of planned cooked quantity and contextual factors, and can suggest an updated cooked quantity to keep leftovers below about 10 kilograms.

A. Success Criteria

- **Accuracy:** MAE below naive baselines and 8–12 kg for a 400-kg/day kitchen.
- **Stability:** Consistent outputs by day and weather conditions.
- **Interpretability:** Model explanations readily understandable by non-technical users.

IV. Data and Assumptions

A synthetic 500-day data set of GIM's eating patterns was used, each record a combination of lunch and dinner.

A. Key Fields

- **date, day of week, weekend flag**
- **event type**
- **weather features** (temperature, rainfall, humidity)
- **main courses, dessert type**
- **total mass cooked, estimated consumption, resulting wastage**

B. Assumptions

- Weather conditions approximate usual Goan monsoons.
- Festivals and exams are of regular campus frequency.
- Some desserts stimulate appetite and affect consumption.
- No count of attendance is observed; total amount cooked is used as a proxy for expected turnout.

V. Exploratory Insights

Exploration showed that waste will rise if there is more food cooking. Weekends and holidays influence consumption, rain reduces turnout, and yesterday's waste provides excellent prediction data. The above observations render it rightful to include lag and calendar features in the model.

VI. Model Selection Rationale

A Decision Tree Regressor was employed because of its readability, effectiveness, and simplicity to interpret. It generates simple-to-interpret if–then rules (e.g., "if cooked quantity is above 410 and a holiday, expected waste increases").

This makes the model useful for non-technically trained staff members.

More sophisticated ensemble methods such as Random Forests or XGBoost might improve slightly in accuracy but would have to be more set up and would be harder to interpret. The Decision Tree is a good, clear baseline before exploring such methods in follow-up work.

VII. Methodology

A. Data Preparation

Records were date sorted, duplicates removed, and categorical variables such as event type and dessert category converted to binary indicators. A lag variable for yesterday's waste was added.

B. Feature Set

Input variables were cooked quantity, temperature, rainfall, humidity, weekend indicator, event and dessert flags, and yesterday's waste. Target variable was food waste in kilograms.

C. Training and Validation

Data was split temporally into 70% training, 15% validation, and 15% test sets to avoid leakage. Hyperparameters were tuned on the validation set, and the optimal model was retrained on combined training and validation data.

D. Baselines

Two trivial predictors were used for comparison: (1) yesterday's waste, and (2) a three-day moving average. Forecast skill was estimated with respect to these baselines using MAE.

E. Model Configuration

A small grid search was performed over tree depth (3–7) and minimum samples per leaf (1–5).

F. Metrics

MAE captured average daily error, RMSE penalized larger errors, R^2 provided quality of fit, and coverage captured frequency of prediction within a 10-kg bound.

VIII. Results and Discussion

The best configuration (tree depth = 5, min leaf size = 1) achieved MAE = 6.40 kg, RMSE = 8.10 kg, and $R^2 = 0.786$ on the test set.

A. Interpretation

An MAE of 6.4 kg represents about 1.6% of mean daily production, within a reasonable range for operations. RMSE values that are close to MAE mean that the model only made a few large mistakes. The high R^2 value means that the model accounted for most of the waste variance.

B. Drivers

The best predictors were quantity cooked and yesterday's waste, with event and weather indicators in second place. These results reflect common sense—planned volume and recent history are strong predictors of leftovers.

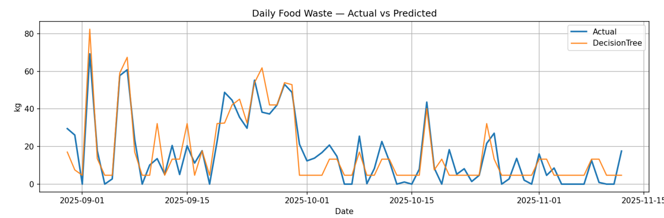

Fig. 1. Actual vs. predicted food waste on the test period.

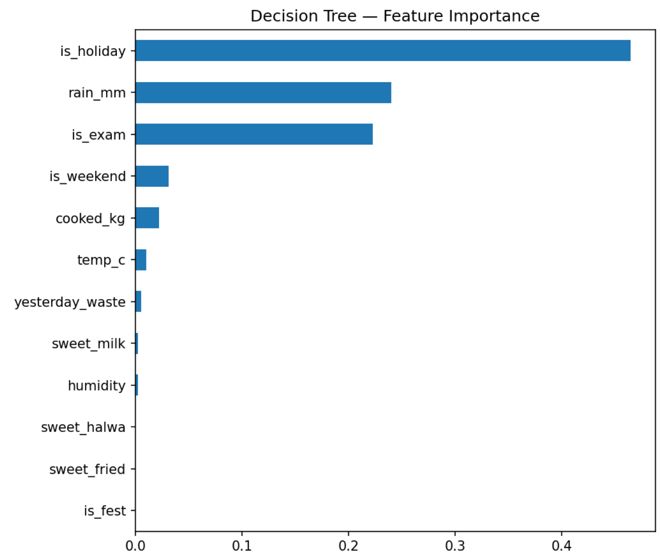

Fig. 2. Feature importance ranking of the Decision Tree model.

TABLE I
Model Performance on Test Data

Model	MAE (kg)	RMSE (kg)	R^2
Decision Tree (depth = 5, leaf = 1)	**6.40**	**8.10**	**0.786**

IX. Operational Integration

In practice, the model would be run nightly before cooking. Staff can enter amounts of planned cooking, weather information, and events. The model works out expected waste and recommends an adjusted amount of cooking that has a buffer for safe leftover, typically 10 kilograms. Everyday planning is more factual without added complexity.

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

We evaluate models with an **expanding (rolling-origin) backtest** using horizon $H{=}1$ day. At each time t, we fit on all data $\{1, \ldots, t\}$ and predict $t{+}1$. As a transparent baseline, we use a **seasonal-naive** predictor

$$\hat{y}_{t+1}^{\text{naive}} = y_{t+1-7},$$

which captures weekly seasonality. Performance is reported as **MAE** and **SMAPE**:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \qquad \mathrm{SMAPE} = \frac{100}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{(|y_i| + |\hat{y}_i|)/2}.$$

We also break out results by calendar month to expose regime effects.

B. Decision Layer: Quantile Regression & Intervals

To support buffer selection we fit quantile models (e.g., $\tau \in \{0.1, 0.5, 0.9\}$) via gradient boosting with the *pinball loss*:

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

This yields median $\hat{q}_{0.5}$ and a 90% predictive interval $[\hat{q}_{0.1}, \hat{q}_{0.9}]$. We report **empirical coverage**:

$$\mathrm{Cov}_{90} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}\{\, y_i \in [\hat{q}_{0.1,i}, \hat{q}_{0.9,i}]\,\}.$$

Buffer policy: kitchen-prepared quantity is

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

trading off shortage vs. waste. We simulate expected *waste* $(\tilde{y} - y)_+$ and *shortage* $(y - \tilde{y})_+$ for $\lambda \in [0,1]$.

C. Event Effects with Hierarchical Pooling

We analyze residuals $r_i = y_i - \hat{y}_i$ by *event type* $e \in \mathcal{E}$ (e.g., exams, holidays). For each event, we estimate a pooled adjustment via empirical Bayes shrinkage:

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

where $\bar{r}_e$ is the mean residual for event e, n_e is its count, $\bar{r}$ is the global residual mean, and k is a prior strength (we use the median n_e). This yields stable effects and a fallback for unseen/rare events.

D. Explainability & Actionable Levers

We report feature influence using **SHAP** (TreeExplainer) when available; otherwise **permutation importance** on the preprocessed design matrix. This surfaces actionable levers (e.g., weekday, lead-time, signals) that most affect forecasted waste.

E. Imputation & Cold-Start Strategy

We impute missing numerics with grouped statistics aligned to operational cycles (e.g., weekday means) and set missing categoricals to `Unknown`. For cold-start or sparse segments, we rely on (i) temporal features (DOW/month/rolling means) and (ii) hierarchical pooling from the event layer to prevent overfitting and provide reasonable priors.

TABLE II
ROLLING-ORIGIN (H=1) OVERALL 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 III
TOP MISPREDICTIONS WITH CONTEXT (EXCERPT).

Date	Actual (kg)	Pred (kg)	—Error— (kg)	DOW	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

F. Error Analysis: Top Mispredictions

We list the top-k absolute errors with their contextual features to support root-cause review (e.g., unusual events, supply shocks, occupancy swings).

X. MODEL MAINTENANCE

- Keep taking data on a daily basis to retrain.
- Retrain every week or month when new data are accumulated.
- Monitor rolling MAE and RMSE and notify unusual shifts.
- Document menu or scheduling revisions likely to affect demand patterns.

XI. LIMITATIONS AND ETHICAL CONSIDERATIONS

The model is currently derived from synthetic data; actual-day records would be required for deployment. Weather inputs are simplified, and intra-day variation is not represented. Ethical use includes keeping the model running to avoid waste, rather than cutting portions disproportionately. No personal data is collected or stored.

XII. FUTURE WORK

- Utilize real mess data and other lag features.
- Include student engagement measures and best-requiring dishes.
- Attempt ensemble models for incremental accuracy improvement.
- Develop a light-weight dashboard for scenario analysis and cook-quantity recommendations.

XIII. REPRODUCIBILITY STATEMENT

All code, figures, and notebooks are public on GitHub. The Colab notebook replicates all reported metrics and plots here.

XIV. AI ASSISTANCE DECLARATION

Software for language generation was used only for preparation and formatting of support. Authors conducted and validated all data analysis, interpretation, and revisions.

XV. Conclusion

This study gives a realist, interpretable baseline to estimate daily food wastage in a campus mess. The Decision Tree model is at the right balance between simplicity, interpretability, and accuracy, giving actionable insights towards operational planning. With incremental rollout of actual data and refinement over time, the framework promises to contribute meaningfully towards sustainability and resource efficiency at GIM and similar institutions.

Acknowledgment

The authors thank the GIM mess staff and the faculty mentors for their guidance and operating experience during this endeavor.

References

[1] D. J. Hand, "Classifier technology and the illusion of progress," *Statistical Science*, 21(1), 2006, pp. 1–14.

[2] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of Statistics*, 29(5), 2001.