Model Development Phase Template

Date	8 Dec 2025
Team ID	RT
Project Title	Restaurant Recommendation System
Maximum Marks	5 Marks

Feature Selection Report Template

In the forthcoming update, each feature will be accompanied by a brief description. Users will indicate whether it's selected or not, providing reasoning for their decision. This process will streamline decision-making and enhance transparency in feature selection.

Feature	Description	Selected (Yes/No)	Reasoning
cuisines	Comma-separated cuisines (e.g., North Indian, Biryani)	Yes	Core for TF-IDF "soup" vectorization; 2723 unique combos drive similarity matching.
name	Restaurant name (e.g., Domino's Pizza)	Yes	Fuzzy/partial matching boosts chain recs (e.g., Domino's → other locations).
location	Bangalore area (e.g., BTM, HSR)	Yes	Location proximity essential for practical recs; top 93 areas clustered.
rate	Avg rating (0-5 scale, mean 4.04)	Yes	Sorts top-K recs by quality; post-fill median reliable predictor.
approx_cost(for two people)	Cost in INR (~800 mean)	Yes	Filters by budget; numeric after cleaning, correlates weakly with rate.
online_order	Boolean (Yes/No)	No	Binary, low variance (not cuisine-specific); irrelevant for content recs.
book_table	Boolean table booking	No	Service feature, sparse data (~10%); doesn't aid taste matching.
votes	Total votes (mean ~100)	No	Popularity proxy but skewed; rate sufficient for quality.
rest_type	Type (Casual, Fine Dining)	No	Overlaps cuisines; redundant for TF-IDF soup.

dish_liked	Top dishes (54% missing)	No	High missing rate; too sparse for vectorization.
reviews_list	Raw review texts	No	Unstructured/noisy; too compute-heavy vs cuisine gain.
menu_item	Specific items (sparse)	No	Low fill rate; covered by cuisines.
listed_in(type/city)	Zomato categories	No	Metadata duplicates location/cuisines; low unique value.