Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

An Early-Warning Risk Signal System for Food Safety Enforcement

Jayshree Ghorpade-Aher
DCET, SoCSE
MIT WPU Pune, India

Aastha Shinde
DCET, SoCSE
MIT WPU Pune, India

Ashvala Anuj
DCET, SoCSE
MIT WPU Pune, India

Anita Gunjal
DCET, SoCSE
MIT WPU Pune, India

Abstract **- Food safety inspections are limited in terms of resources, meaning that only a small proportion of establishments can be inspected. This means that inspections rarely occur based on early indicators. Customer reviews from sites such as Google Maps often include details of hygiene issues including pests, spoilage, temperature and unhygienic conditions. We propose an advisory early-warning system to help identify high-priority establishments to inspect based on customer reviews. The system includes six steps: collecting reviews, identifying hazards in five hygiene categories (foodborne illness, pest control, cross-contamination, food spoilage, and sanitation and premises), weighting reviews based on credibility, applying time decay, adjusting based on region and calculating the final score. In the current prototype, hazards are identified using keywords rather than a trained classifier, and negation handling is not yet implemented. The final 0 to 100 score and the spatial clustering have not been calibrated or validated. Therefore this system should be considered a prototype scoring and prioritisation framework, not a validated predictor of food safety violations. The score is used only to rank establishments, and all cases need to be verified by an inspector.**

Index Terms **- food safety, early warning, review mining, risk scoring, public health informatics, responsible AI**

I. INTRODUCTION

Food safety enforcement in India still depends largely on manual sweeps and rotational inspections. During a 100-day drive concluding in September 2026, the Maharashtra Food and Drug Administration inspected over 12,000 establishments, issued over 5,200 improvement notices and suspended 603 licences. Such results show the value of enforcement but also the manpower it consumes. At the same time, thousands of hygiene signals are posted daily as public reviews on platforms such as Google Maps, and authorities cannot use them because the text is unstructured, code-mixed, geospatially inconsistent and exposed to fake reviews.

This work proposes a context-aware, explainable pipeline that turns noisy public feedback into ranked, location-precise risk scores to support inspector deployment. Its objectives are: (i) a spatial pipeline that maps reviews by latitude and longitude rather than text addresses; (ii) a Transformer-based NLP stage that separates genuine hazards from slang and negation; (iii) a time-decay model so that recent complaints outrank old ones; and (iv) a unified 0 to 100 score that combines text evidence, credibility, space and time. The scope is digitally visible food establishments in Pune with a Google Maps review presence. The score is a ranking signal and not a calibrated probability of non-compliance, because no FSSAI inspection outcome data was available for validation.

II. LITERATURE SURVEY

Recent research highlights the potential of online consumer reviews and machine learning for improving food safety surveillance. The study on passive surveillance of foodborne illnesses in Indian restaurants [1] applied BERT-based aspect-based sentiment analysis (ABSA) and hygiene keyword detection to identify food safety concerns, while highlighting challenges related to review reliability and regulatory verification. Similarly, research on food safety risk detection [2] employed natural language processing and machine learning classifiers to identify hygiene risks and support preventive measures.

Kohli and Garg [3] examined India's food safety challenges, emphasizing inadequate surveillance, underreporting, and regulatory limitations. The Eat at Right Place initiative [4] demonstrated the effectiveness of BERT-based sentiment analysis in extracting restaurant-related concerns from consumer reviews. Siering [5] explored the prediction of restaurant health violations using online reviews, demonstrating their potential for supporting inspection decisions.

Furthermore, domain-specific food safety prescreening methods [6] and social media surveillance approaches [7] highlighted the usefulness of consumer-generated data for identifying potential risks. Beyond food safety, Ghorpade and Sonkamble [8] combined ensemble learners (gradient boosting, random forest and XGBoost) with weighted probabilities and SHAP-based model explanations to generate predictive health-risk recommendations, showing the value of explainable risk prediction for public health decision support. Despite these developments, existing approaches face limitations involving data quality, geographical applicability, language diversity, and verification of reported incidents. These gaps motivate the development of an integrated early-warning system that identifies emerging hygiene concerns, monitors risk patterns, and supports evidence-based food safety enforcement.

III. PROPOSED SYSTEM

This section describes the full proposed design; the components implemented in the current prototype are reported in Section IV. The system is a sequential decision-support pipeline of six engines (Fig. 1). Reviews are collected from Google Maps using Playwright, because the platform loads content through JavaScript, and cleaned with Pandas (duplicate removal, missing-value checks, type conversion). Each record holds the restaurant name, review text, rating, review date, map URL, scrape timestamp, review identifier and source platform.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

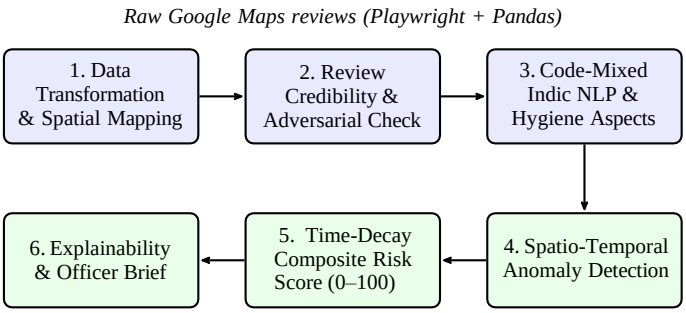


Fig. 1. Six-engine architecture of the proposed system.

A. Data Transformation and Spatial Mapping

Records are validated and given a data-quality status (valid, partially usable, spatially incomplete, temporally uncertain or rejected) instead of being deleted, so a bad coordinate does not remove a useful review from the NLP stage. Reviews are linked to an establishment through a stable place identifier. Latitude and longitude are extracted directly from the map URL, since Indian addresses suffer from spelling variation, missing building numbers and informal area names, and a Maharashtra boundary check catches wrong coordinates. Relative dates such as "3 months ago" are converted to an estimated absolute date using the scrape timestamp, with an uncertainty interval so that an estimate is never treated as an exact legal fact.

B. Review Credibility and Adversarial Detection

Each review is embedded with Sentence-Transformers all-MiniLM-L6-v2, and cosine similarity is computed between reviews of the same establishment within a 48-hour window. Highly similar text, identical ratings and close timestamps raise an adversarial suspicion value A_i. Because a genuine outbreak also produces similar wording, a separate organic-outbreak evidence value is kept. Further features are account history, text detail, rating-text consistency, copy similarity and reviewer diversity. Reviews are down-weighted and never deleted:

$$C_{w,i} = C_{\text{base},i}(1 - \rho A_i) \tag{1}$$

where ρ limits the maximum down-weighting. A single-source safeguard labels a lone complaint as "requires verification" and prevents it from creating the highest priority on its own.

C. Code-Mixed Indic NLP and Hygiene Aspect Extraction

Whole-review sentiment can classify "the ambience was beautiful, but I found a cockroach in the soup" as positive, so the system performs aspect-based hygiene analysis. Reviews mix English, Hindi, Marathi and transliterated text (for example "khana baasi tha"), which English-only BERT handles poorly; IndicBERT is therefore proposed for primary inference, with MuRIL as a comparison model. In the proposed design, negation words (no, not, nahi) are preserved, so "no cockroaches" is not scored as a pest complaint. Each sentence receives a hygiene probability, and hygiene sentences are converted to aspect-condition pairs, for example (soup, pest observed). Evidence is mapped to the categories in Table I (weights shown in Fig. 2) and more than one category may be retained per review. The NLP contribution is $S_{\text{NLP},i} = S_{\text{cat},i} \times P_i \times Q_i$, where P_i is model confidence and Q_i is extraction quality.

TABLE I
HYGIENE CATEGORIES AND SEVERITY WEIGHTS

Category	Example signals	Severity
Foodborne illness	vomiting, diarrhoea, food poisoning	1.00
Pest control	cockroach, rat, insect, fly	0.85
Cross-contamination	raw chicken, dirty hands, unwashed	0.80
Food spoilage	stale, rotten, undercooked	0.70
Sanitation and premises	dirty, unclean, washroom, waste	0.55

Severity values are design parameters, not legal conclusions.

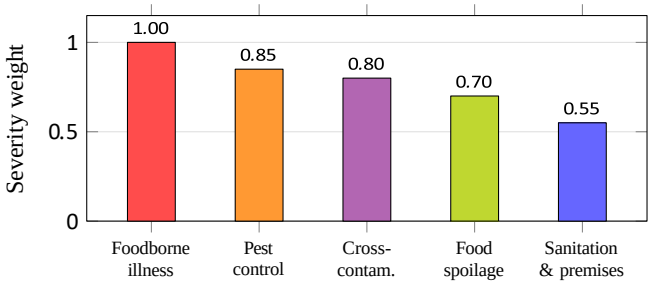


Fig. 2. Severity weights of the five hygiene categories.

D. Spatio-Temporal Anomaly Detection

Similar complaints from nearby establishments in a short period may indicate a shared cause such as contaminated water or ice, a common supplier or a pest infestation. ST-DBSCAN is applied with Haversine distance as the spatial measure. Two signals are neighbours only if their distance is at most 400 m and their time difference is at most 72 hours, and a cluster needs at least three distinct establishments, so three reviews from one restaurant are not a regional hazard. A category-compatibility rule keeps unrelated complaints, such as slow service, out of clusters. The regional multiplier M_{reg} is 1.5 for an active high-confidence cluster, 1.2 for a weak one and 1.0 for an isolated signal. It is capped, and the result is described as a possible regional hazard and not proof of contamination. The 400 m and 72 h values are starting parameters for sensitivity analysis.

E. Time-Decay Composite Risk Scoring

The contribution of review i combines category severity $S_{\text{cat},i}$, NLP confidence P_i, credibility weight $C_{w,i}$ and evidence quality Q_i:

$$S_i = S_{\text{cat},i} \times P_i \times C_{w,i} \times Q_i \tag{2}$$

Older reviews are discounted with an exponential half-life function, where t_i is the review age in days and H is the half-life:

$$W(t_i) = 2^{-t_i/H} \tag{3}$$

With $H = 90$ days the weight is 0.794 at 30 days, 0.500 at 90 days, 0.250 at 180 days and 0.060 at 365 days (Fig. 3). The establishment base risk sums the decayed contributions of

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

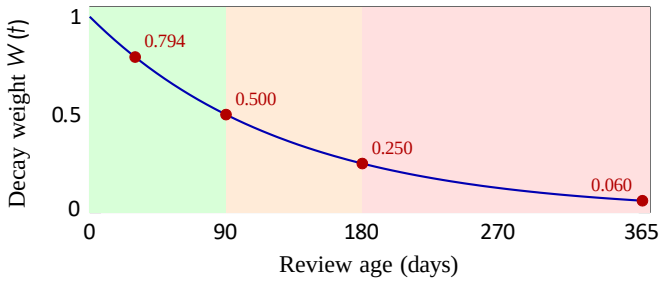


Fig. 3. Time-decay weight of a review with half-life $H = 90$ days.

its reviews E_j. It is multiplied by the regional multiplier of Section III-D and then mapped to a bounded score:

$$R_{\text{base},j} = \sum_{i \in E_j} S_i \, W(t_i) \tag{4}$$

$$R_{\text{regional},j} = R_{\text{base},j} \times M_{\text{reg},j} \tag{5}$$

$$R_{\text{final},j} = 100 \left(1 - e^{-\alpha R_{\text{regional},j}}\right) \tag{6}$$

where α is a scaling constant that must be calibrated. The score cannot exceed 100 and rises gradually with evidence. Operational bands are Low (0 to 29), Medium (30 to 59), High (60 to 79) and Critical (80 to 100). Safeguards include a minimum independent-review rule, a cap on repeated text, category diversity, temporal freshness and mandatory human verification, so a high score triggers an inspection recommendation and never an automatic penalty.

F. Explainability and Officer Brief Generation Engine

For establishments above a threshold (for example 60) the system generates an inspection brief with the establishment, location, score, band and rank. The model explanation shows why the classifier assigned a category (for example Integrated Gradients and highlighted evidence spans), following the use of model explanations to make health-risk predictions interpretable [8], and the score explanation shows how the 0 to 100 value was computed. Every component is traceable to the review identifier, text, source URL, date, model version, category, confidence, credibility weight, decay factor and regional multiplier, giving an audit trail that the officer can verify. Outputs are JSON, HTML or PDF, dashboard cards and CSV.

IV. RESULTS AND DISCUSSION

A. Dataset and Prototype Configuration

The prototype processed 18,831 Google Maps reviews of Pune restaurants. Text was lower-cased, newlines were removed and whitespace was normalised. Hazards were detected with predefined keywords for the five categories of Table I, and reviews with no hygiene keyword received zero severity. Negation handling was not applied in this prototype. The hazard probability is a manually specified value, credibility is a heuristic based on text length, and the half-life is 90 days. The results therefore demonstrate the operation of the scoring framework and not classification accuracy.

B. Restaurant-Level Risk

Review contributions were decayed and summed per restaurant. Base risk differed by about ten times across the sampled establishments, which shows that the framework turns individual review observations into a comparable establishment-level signal. Establishments are anonymised here because the values are uncalibrated prototype outputs.

C. Discussion and Limitations

Temporal weighting keeps the signal current: a 90-day-old review counts half as much as a new one and a 150-day-old review about 31.5%. Aggregation avoids treating a single review as proof of a violation. The main limitations are: the detector is keyword-based and not a trained classifier; negation is not handled; the probability is manually assigned; credibility is a heuristic and not a measure of authenticity; one dominant category is assigned even when several apply; relative dates are approximate; and spatial clustering and the final 0 to 100 score have not yet been executed or validated. Without inspection outcome data the scores cannot be calibrated, and qualified personnel must verify any case before enforcement action.

V. CONCLUSION AND FUTURE SCOPE

This paper presented an early-warning system that converts public Google Maps reviews into prioritised food safety risk signals through preprocessing, hazard detection, severity weighting, credibility estimation, time decay and restaurant-level aggregation. A prototype on 18,831 reviews showed that the approach produces interpretable establishment-level base risk values. The system is an advisory layer for inspectors and does not replace statutory inspection.

Future work includes (i) replacing keyword rules with a fine-tuned IndicBERT or MuRIL classifier trained on annotated code-mixed reviews, including negation handling; (ii) integrating ST-DBSCAN clustering, calibrating α and H, and validating scores against FSSAI inspection outcomes; (iii) a real-time pipeline that continuously collects new reviews and updates risk signals; and (iv) covering unlicensed vendors and street stalls through municipal registries or citizen reporting.

REFERENCES

[1] "Utilising consumer reviews for passive surveillance of foodborne illnesses: Insights and challenges from the Indian restaurant," *International Journal of Public Health Science (IJPHS)*, vol. 14, pp. 479–492, Mar. 2025.
[2] "From detection to resolution: Using machine learning to identify food safety risks and solutions in online consumer reviews," *European Journal of Marketing*, 2024.
[3] C. Kohli and S. Garg, "Food safety in India: An unfinished agenda," 2015.
[4] "Agile fusion: Developing 'Eat at Right Place' sentiment analysis tool," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 34, pp. 602–619, Apr. 2024.
[5] M. Siering, "Leveraging online review platforms to support public policy: Predicting restaurant health violations based on online reviews," *Decision Support Systems*, 2021.
[6] "A food safety prescreening method with domain-specific information using online reviews," *Journal of Consumer Protection and Food Safety*, vol. 17, pp. 163–175, 2022.
[7] "A research framework for passive surveillance for food safety from social media: Identification and evaluation of customer reviews for regulatory use and case study of 30 restaurants," 2022.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

[8] J. Ghorpade and B. Sonkamble, "Predictive recommendations for diabetes using ensemble techniques with model explanations," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 11, no. 4, pp. 413–421, 2023. [Online]. Available: https://ijisae.org/index.php/IJISAE/article/view/3538

(This work is licensed under a Creative Commons Attribution 4.0 International License.)