Proceedings of the 4th International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/2026.36860

Dynamic Identification of Hotel Service Experience Quality Based on Semantic Mining of Online Reviews

Weijia Liu

College of Business, Law and Social Sciences, University of Derby, Montreux, Switzerland
rara481846778@gmail.com

Abstract. This study develops a dynamic framework for identifying hotel service experience quality through semantic mining of online reviews. Ctrip reviews from 260 hotels in Shanghai, Beijing, Guangzhou, Chengdu, and Hangzhou between 2021 and 2024 were cleaned into 156,000 valid reviews and 10,842 hotel-month observations. The framework combines a service-experience ontology, dependency-based aspect extraction, SKEP sentiment scoring, complaint intensity, change-point detection, and Gaussian hidden Markov modeling across eight service dimensions. Results show that semantic indicators detect deterioration earlier than numerical ratings, especially in cleanliness, facility reliability, and problem recovery, where complaint intensity rose 18%–27% above the sample mean during risk periods. The model classifies four quality states and achieves an F1-score of 0.829 and AUC of 0.903 for predicting next-month service decline, providing a reproducible method for dynamic hotel quality monitoring.

Keywords: Hotel service experience, online reviews, semantic mining, dynamic identification, quality change detection

1. Introduction

Hotel service quality can no longer be fully represented by star ratings or overall satisfaction scores. Online reviews contain detailed information on rooms, staff, hygiene, facilities, catering, location, value, and service recovery, often revealing mixed opinions across multiple dimensions within the same review [1]. Hospitality sentiment analysis has therefore moved from overall polarity toward aspect-level semantic analysis, with performance influenced by preprocessing, feature representation, language context, and domain-specific aspect extraction [2].

Review text also supports dynamic monitoring of hotel performance [3]. Tracking sentiment and complaint patterns over time can reveal service decline before overall ratings change substantially, especially as guest expectations vary across crisis, recovery, and market conditions [4, 5].

This study develops a semantic mining framework that converts reviews into hotel-month indicators, detects shifts in complaint intensity and sentiment, and classifies monthly quality states using a Gaussian hidden Markov model. By combining aspect-level semantics, rating deviation, reply behavior, and hidden-state transitions, the framework provides early warning signals for hotel service deterioration.

DOI: 10.54254/2755-2721/2026.36860

2. Literature review

2.1. Online reviews and hotel service experience

Online hotel reviews provide direct accounts of guest experiences, covering service, facilities, rooms, dining, and service recovery. They reveal both tangible and intangible quality dimensions, including cleanliness, reliability, responsiveness, empathy, ambience, and process quality [6]. Compared with questionnaires, reviews are spontaneous, time-stamped, and tied to actual consumption, making them suitable for fine-grained temporal analysis. However, noise, emotional exaggeration, repetition, and platform-specific language require semantic cleaning and dimension mapping before evaluation.

2.2. Aspect-level semantic mining of service quality

Aspect-level semantic mining links service objects with emotional evaluations across predefined hotel service dimensions. This is important because a single review may express different sentiments toward multiple aspects. Prior research shows that aspect-based sentiment analysis can identify service attributes, negative experiences, and service failures [7, 8], while deep learning can convert reviews into measurable improvement indicators [9]. Compared with document-level sentiment, aspect-level analysis provides clearer operational diagnosis by identifying the specific service dimension responsible for satisfaction or complaint.

2.3. Dynamic identification of service experience change

Hotel service quality changes over time due to staff turnover, renovation, seasonality, travel demand, and platform visibility. Static averages can hide these shifts and delay managerial response. Prior studies therefore emphasize tracking temporal sentiment and service-attribute trajectories, while evidence also shows that seasonality, travel purpose, and customer background affect satisfaction patterns [10]. Dynamic identification requires monthly semantic indicators, change-point detection, and interpretable quality states to determine whether a hotel is stable, improving, declining, or entering a service risk-warning stage.

3. Methodology

3.1. Review dataset and hotel-month observation construction

Ctrip reviews from 260 hotels in Shanghai, Beijing, Guangzhou, Chengdu, and Hangzhou between January 2021 and December 2024 were collected using Scrapy 2.11 and stored in MySQL 8.0. Records included hotel information, review text, ratings, dates, travel type, helpful counts, and reply status.

After removing short, duplicate, incomplete, invalid, and template-based reviews, 156,000 valid samples remained. Reviews were aggregated into hotel-month units, and months with fewer than 12 reviews were excluded to reduce instability in dynamic modeling.

The hotel-month semantic observation was defined as shown in Equation (1) and (2):

$$Xh,t = [\phi(Rh,t), \mu(S_{h,t}), \rho^{-}(h,t), \kappa(h,t), \Delta q_{h,t}, \eta_{h,t}, \omega_{h,t}, \pi_{h,t}] \tag{1}$$

DOI: 10.54254/2755-2721/2026.36860

$$R_{h,t} = r_i | hotel_i = h, month_i = t \tag{2}$$

where $\phi(Rh,t)$ denotes dimension mention structure, $\mu(S_{h,t})$ denotes average semantic sentiment intensity, $\rho^{-}(h,t)$ denotes negative sentiment ratio, $\kappa(h,t)$ denotes complaint intensity, $\Delta q_{h,t}$ denotes rating deviation from the previous three-month mean, $\eta_{h,t}$ denotes helpful complaint ratio, $\omega_{h,t}$ denotes hotel reply coverage, and $\pi_{h,t}$ denotes detected change-point status. This structure allows each month to be treated as a comparable semantic quality unit.

3.2. Service-experience ontology and aspect-sentiment extraction

A fixed hotel service ontology was developed across eight dimensions: room environment, cleanliness, staff service, facility reliability, catering, location, price value, and problem recovery. Python 3.11 and HanLP 2.1 were used for segmentation, part-of-speech tagging, and dependency parsing. Aspect terms were extracted through dependency patterns linking service objects with adjectives, verbs, negations, and degree adverbs, capturing expressions such as "room clean," "wait too long," and "refund delayed."

Dimension assignment followed dictionary matching and dependency-context scoring as shown in Equation (3):

$$d(a_j) = \arg\max_{d \in D} [\lambda_1 M(a_j, L_d) + \lambda_2 C(a_j, d) + \lambda_3 G(a_j, d) - \lambda_4 N(a_j, d)] \tag{3}$$

where a_j is the extracted aspect term, D is the ontology set, $M(a_j, L_d)$ measures lexical matching between the aspect and dimension lexicon, $C(a_j, d)$ measures contextual co-occurrence, $G(a_j, d)$ measures dependency-grammar compatibility, and $N(a_j, d)$ penalizes ambiguous or negated mappings. The parameters were fixed as $\lambda_1 = 0.40, \lambda_2 = 0.25, \lambda_3 = 0.25,$ and $\lambda_4 = 0.10.$

Sentiment polarity and strength were calculated with the Baidu SKEP Chinese sentiment model. Each aspect sentence received positive probability, negative probability, and semantic intensity. Complaint intensity was computed from negative aspect sentences containing hygiene, delay, noise, equipment failure, refund, attitude, and compensation terms as shown in Equation (4):

$$C_{h,t,d} = \frac{1}{n_{h,t,d}} \sum_{j=1}^{n_{h,t,d}} \left[p_j^{-} \cdot \left(1 + \alpha g_j + \beta n_j + \gamma u_j \right) \right] \cdot w_d \tag{4}$$

where p_j^{-} is negative sentiment probability, g_j is degree-adverb strength, n_j is negation conflict, u_j is urgency vocabulary, and w_d is the dimension weight. This formula gives stronger weight to severe operational complaints than to mild negative descriptions.

3.3. Dynamic quality indicator and temporal state model

For each hotel-month, eight indicators were calculated, including mention ratio, sentiment ratios and intensity, complaint intensity, rating deviation, helpful complaint ratio, and hotel reply coverage. After MinMax normalization, these variables formed monthly semantic quality vectors. Structural

DOI: 10.54254/2755-2721/2026.36860

shifts in complaint intensity and sentiment decline were detected using Bayesian change-point analysis with the PELT algorithm and rbf cost function. A Gaussian hidden Markov model then classified monthly service quality into four states: stable high quality, gradual improvement, gradual decline, and service risk warning.

4. Experimental design

4.1. Data cleaning, semantic parsing, and monthly indicator generation

Raw review records were cleaned by removing duplicated text, platform boilerplate, emojis, excessive punctuation, hotel address fragments, and merchant template replies. Chinese text was segmented with HanLP, and dependency parsing was applied at the sentence level. Extracted aspect terms were mapped to the eight ontology dimensions. SKEP then returned sentiment probability for every aspect sentence. Monthly indicators were calculated for each hotel and each dimension. The final panel contained hotel ID, city, star level, month, semantic indicators, rating statistics, complaint indicators, reply indicators, and quality-state labels.

4.2. Change-point detection and hidden-state identification workflow

The monthly semantic sequence of each hotel was arranged chronologically. The ruptures algorithm applied PELT search with an rbf cost function to identify structural shifts in sentiment decline and complaint intensity. Detected change points were converted into binary monthly flags and merged into the HMM input matrix. The Gaussian HMM was trained with n_components = 4, covariance_type = diagonal, n_iter = 200, and random_state = 42. After training, each hotel-month observation was assigned to one dynamic service-quality state.

4.3. Validation, interpretation, and output generation

Validation used three outputs: consistency with user rating movement, warning accuracy for next-month rating decline, and complaint-surge detection accuracy. A hotel-month was marked as a confirmed risk case when the next-month average rating decreased by more than 0.4 points and the negative sentiment ratio increased by more than 15%. HMM warning results were compared with confirmed risk labels using precision, recall, F1-score, and AUC. Dimension-level interpretation was generated by ranking standardized indicator deviation from the hotel's previous three-month mean.

5. Results

5.1. Descriptive statistics of review texts, hotel attributes, and service-experience dimensions

The final panel retained 10,842 hotel-month observations. Each hotel-month contained 14.38±5.72 valid reviews after sparse-month filtering. Aspect extraction produced 418,673 valid aspect sentences, with 3.07±1.26 aspects per review. Cleanliness, room environment, and staff service were the most frequently mentioned dimensions, while problem recovery showed the strongest negative deviation (see Table 1).

DOI: 10.54254/2755-2721/2026.36860

Table 1. Descriptive statistics of hotel service-experience dimensions

Service dimension	Mention intensity	Positive semantic score	Negative semantic score	Complaint intensity	Three-month deviation
Room environment	0.182±0.041	0.716±0.083	0.194±0.052	0.231±0.064	-0.118±0.037*
Cleanliness	0.169±0.038	0.682±0.091	0.238±0.061	0.284±0.073	-0.164±0.042**
Staff service	0.157±0.035	0.741±0.079	0.176±0.047	0.205±0.059	-0.092±0.031*
Facility reliability	0.136±0.032	0.651±0.088	0.261±0.066	0.319±0.081	-0.181±0.049**
Catering experience	0.104±0.027	0.694±0.085	0.214±0.056	0.247±0.069	-0.126±0.040*
Location convenience	0.098±0.025	0.782±0.071	0.143±0.039	0.152±0.044	0.046±0.018
Price value	0.081±0.022	0.663±0.087	0.229±0.058	0.269±0.071	-0.137±0.044*
Problem recovery	0.073±0.019	0.604±0.096	0.312±0.074	0.386±0.093	-0.218±0.057**

*Note: *p<0.05; **p<0.01 compared with the stable high-quality state.*

5.2. Change-point detection and dynamic service-quality state identification results

The PELT-HMM workflow detected 1,284 structural change points. Service risk-warning months had higher complaint intensity, stronger negative sentiment, and larger rating deviation than gradual-decline months. The transition matrix showed that stable high-quality months had strong persistence, while gradual-decline months were most likely to move into risk warning when complaint intensity exceeded the previous three-month mean by 0.20±0.06 (see Figure 1).

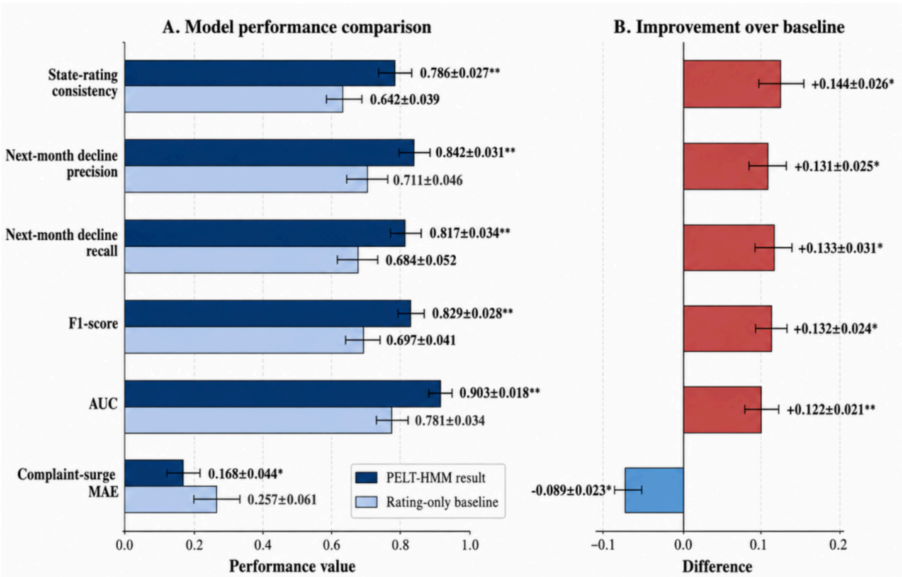

Figure 1. Performance of dynamic service-quality state identification

DOI: 10.54254/2755-2721/2026.36860

5.3. Dimension-level risk diagnosis and warning performance for hotel service experience quality

Dimension-level diagnosis showed that cleanliness, facility reliability, and problem recovery contributed most to warning-state formation. In confirmed risk months, problem recovery had the largest standardized deviation, followed by facility reliability and cleanliness. These results indicate that guests do not move into strong dissatisfaction only because of isolated negative emotions. Risk emerges when operational complaints, weak recovery signals, and rating deviation accumulate across consecutive months. The HMM warning state therefore provides earlier and more interpretable service-quality signals than average rating decline alone.

6. Conclusion

This study develops a dynamic semantic mining framework for hotel service quality using online reviews. It converts review text into hotel-month indicators, detects structural changes, and classifies temporal quality states. Results show that aspect-level sentiment, complaint intensity, and reply coverage can identify deterioration before ratings decline, with cleanliness, facility reliability, and problem recovery being the most sensitive dimensions. The framework supports dynamic monitoring, operational risk detection, and timely service recovery. Future work should incorporate multilingual, cross-platform, and real-time review data.

References

[1] Mehraliyev, F., Chan, I. C. C., & Kirilenko, A. P. (2022). Sentiment analysis in hospitality and tourism: A thematic and methodological review. *International Journal of Contemporary Hospitality Management, 34*(1), 46–77.

[2] Ameur, A., Hamdi, S., & Ben Yahia, S. (2023). Sentiment analysis for hotel reviews: A systematic literature review. *ACM Computing Surveys, 56*(2), Article 51.

[3] Bi, J.-W., Zhu, X.-E., & Han, T.-Y. (2024). Text analysis in tourism and hospitality: A comprehensive review. *Journal of Travel Research, 63*(8), 1847–1869.

[4] Nicolau, J. L., Xiang, Z., & Wang, D. (2024). Daily online review sentiment and hotel performance. *International Journal of Contemporary Hospitality Management, 36*(3), 790–811.

[5] Leoni, V., & Moretti, A. (2024). Customer satisfaction during COVID-19 phases: The case of the Venetian hospitality system. *Current Issues in Tourism, 27*(3), 396–412.

[6] Kalnaovakul, K., & Promsivapallop, P. (2023). Hotel service quality dimensions and attributes: An analysis of online hotel customer reviews. *Tourism and Hospitality Research, 23*(3), 420–440.

[7] Özen, İ. A., & Özgül Katlav, E. (2023). Aspect-based sentiment analysis on online customer reviews: A case study of technology-supported hotels. *Journal of Hospitality and Tourism Technology, 14*(2), 102–120.

[8] Jeong, N., & Lee, J. (2024). An aspect-based review analysis using ChatGPT for the exploration of hotel service failures. *Sustainability, 16*(4), 1640.

[9] Liu, X.-X., & Chen, Z.-Y. (2022). Service quality evaluation and service improvement using online reviews: A framework combining deep learning with a hierarchical service quality model. *Electronic Commerce Research and Applications, 54*, 101174.

[10] Carvalho, F., Ramos, R. F., & Fortes, N. (2024). Customer satisfaction in mountain hotels within UNESCO Global Geoparks: An empirical study based on sentiment analysis of online consumer reviews. *Tourism & Management Studies, 20*(1), 35–47.