Acta Informatica Pragensia
2027, Volume 16, Issue 1, Forthcoming article
https://doi.org/10.18267/j.aip.324

PRAGUE UNIVERSITY OF ECONOMICS AND BUSINESS

Article Open Access

Adaptive Personalization for Tourism Recommendations: A Hybrid CBR-AHP-TOPSIS Framework with Context-Aware Mode Adjustment

Edi Faizal, Thomas Edyson Tarigan, Sumiyatun Sumiyatun

Faculty of Information Technology, Indonesia Digital Technology University, Yogyakarta, Indonesia

Corresponding author: Edi Faizal (edifaizal@utdi.ac.id)

Editorial Record

First submission received:
April 26, 2026

Revisions received:
June 27, 2026
July 5, 2026
July 17, 2026

Accepted for publication:
July 21, 2026

Academic Editor:
Miloslav Hub
University of Pardubice, Czech Republic

This article was accepted for publication by the Academic Editor upon evaluation of the reviewers' comments.

How to cite this article:
Faizal, E., Tarigan, T. E., & Sumiyatun, S. (2027). Adaptive Personalization for Tourism Recommendations: A Hybrid CBR-AHP-TOPSIS Framework with Context-Aware Mode Adjustment. *Acta Informatica Pragensia*, 16(1), Forthcoming article. https://doi.org/10.18267/j.aip.324

Copyright:
© 2027 by the author(s). Licensee Prague University of Economics and Business, Czech Republic. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution License (CC BY 4.0).

Abstract

Background: Conventional case-based reasoning (CBR) recommender systems employ fixed similarity weights, failing to capture dynamic user preferences across travel contexts.

Objective: This paper proposes an adaptive hybrid framework integrating CBR with multi-criteria decision making (AHP and TOPSIS) for context-aware personalization in tourism recommendations.

Methods: Weights of criteria are derived from user surveys using the analytic hierarchy process (AHP) and fine-tuned using transaction-informed mode-specific adjustment factors calculated from historical rating patterns. The framework is evaluated on a real-world tourism dataset from Yogyakarta, Indonesia (15,870 records).

Results: The proposed method achieves a precision of 0.85, recall of 0.88, F1-score of 0.86 and RMSE of 0.27. Compared to feature-enhanced deep learning baselines (LightGCN-feature, SASRec-feature), it achieves 4.9% higher precision and 3.6% higher F1-score. Statistical tests (Friedman, Wilcoxon) confirm medium-to-large effect sizes (r = 0.63–0.92). Cross-regional validation (Bali, Malang), cold-start evaluation and bootstrap confidence intervals support the robustness of the framework within the tested regional subsets, particularly in sparse-data scenarios.

Conclusion: The proposed framework offers an interpretable, generalizable and efficient approach to context-aware personalization, with documented practical implications for destination management organizations.

Index Terms

Recommender systems; Case-based reasoning; Multi-criteria decision making; Analytic hierarchy process; TOPSIS; Context-aware personalization; Tourism informatics.

1 INTRODUCTION

The digital transformation of the tourism industry has led to an explosion of user-generated content and transactional data, creating both opportunities and challenges for personalized service delivery (Raza et al., 2026; Zhang et al., 2025). Travellers increasingly rely on online platforms to plan their trips, but the sheer volume of available information often leads to decision paralysis. Intelligent recommendation systems have emerged as a critical solution to mitigate information overload, yet their effectiveness hinges on the ability to capture the complex, often conflicting criteria that influence human decision-making, such as costs, travel time, comfort, cultural relevance, crowding and facility quality (Flórez et al., 2025; Payandenick et al., 2026).

Traditional recommendation approaches, including collaborative filtering (CF) and content-based methods, suffer from well-documented limitations: cold-start problems, static preference modelling, lack of interpretability and difficulty handling multi-criteria trade-offs (Hassan et al., 2026; Raza et al., 2026). Case-based reasoning (CBR) offers a natural alternative by reusing past experiences to solve new problems, providing case-based explanations and mitigating the cold-start issue (Aamodt & Plaza, 1994; Rani et al., 2021). However, conventional CBR systems typically employ fixed or uniform similarity weights, failing to capture the dynamic and hierarchical nature of user preferences (Alptekin & Büyüközkan, 2011; Faizal et al., 2026). On the other hand, multi-criteria decision making (MCDM) methods such as the analytic hierarchy process (AHP) and the technique for order preference by similarity to ideal solution (TOPSIS) provide structured mechanisms for deriving importance of criteria and ranking alternatives, but they are often applied in isolation without making use of historical case knowledge (Gul & Topcu, 2015; Ristanović et al., 2025; Yang et al., 2025).

Recent research has begun to explore synergies between CBR and MCDM (Alptekin & Büyüközkan, 2011; Chen et al., 2024), but existing work remains limited: either CBR is combined with a single MCDM technique (AHP or TOPSIS separately), or the integration is only conceptual without empirical validation on large real-world datasets (Haq et al., 2025). Moreover, few approaches incorporate context (e.g., visit mode, seasonality) into the weighting mechanism and rigorous statistical validation is often absent. This gap is particularly pronounced in emerging markets, where digital infrastructure gaps and diverse socio-cultural characteristics create unique adoption dynamics (Godolja et al., 2025; Huda et al., 2024).

To address these challenges, the present paper proposes an adaptive hybrid CBR-AHP-TOPSIS framework that integrates AHP for criteria weight elicitation (via user surveys) and TOPSIS for robust ranking, within a unified CBR architecture. Unlike the prior CBR-MCDM integration by Alptekin & Büyüközkan (2011), which remained conceptual, our framework is fully implemented and validated on a large-scale dataset. Compared to CBR+AHP or CBR+TOPSIS separately, our contribution lies in the tight three-way integration where AHP provides interpretable criteria weights, TOPSIS resolves multi-criteria trade-offs and CBR enables case-based explanation and cold-start resilience. The context-aware mode adjustment (Section 3.1) – which dynamically rescales AHP weights using transaction-derived factors per travel mode – is, to our knowledge, novel in the CBR literature. Additionally, we provide the first fair comparison between such hybrid CBR-MCDM systems and feature-enhanced deep learning baselines (LightGCN-feature, SASRec-feature). This allows the system to adapt to different travel contexts without requiring per-user AHP surveys. Additionally, the framework incorporates a learning module that updates the case base with user feedback, closing the CBR loop.

The contributions of this study are fourfold:

- A hybrid recommendation architecture that combines CBR, AHP and TOPSIS with context-aware mode-specific weight adjustment.
- A detailed mathematical formulation of the framework with corrected TOPSIS separation measures.
- A fair empirical evaluation on a large-scale dataset comparing against feature-enhanced deep learning baselines.
- Transparent reporting of mode-specific adjustment factors with statistical validation.

The remainder of this paper is organized as follows. Section 2 reviews related work. Section 3 describes the methodology, including the integrated mathematical foundations, dataset, system architecture, baseline methods and evaluation setup. Section 4 reports the experimental results. Section 5 discusses the findings, practical implications and limitations. Section 6 concludes the paper.

2 LITERATURE REVIEW

2.1 Case-based reasoning in recommendation systems

Case-based reasoning is an artificial intelligence paradigm that solves new problems by retrieving and adapting solutions to similar past problems stored in a case base (Aamodt & Plaza, 1994). The classic CBR cycle comprises four phases: retrieve, reuse, revise and retain (Richter & Weber, 2013). In recommendation contexts, CBR offers advantages over collaborative filtering: it handles unstructured data, provides explanations and mitigates the cold-start problem (Rani et al., 2021).

In tourism, CBR has been applied to destination planning, itinerary recommendation and attraction suggestion. For instance, Rani et al. (2021) developed a CBR-based system for Yogyakarta attractions, achieving high accuracy. Duan et al. (2024) found that CBR strikes the best balance between personalization and data privacy. Alfaro et al. (2024) combined CBR with 360° VR for immersive hospitality experiences. Faizal et al. (2026) proposed a context-aware itinerary model with auto-revise. Despite these successes, conventional CBR implementations typically rely on fixed or uniform weights for similarity computation, motivating the need for adaptive weighting.

2.2 Multi-criteria decision making for personalization

MCDM encompasses methods for evaluating alternatives based on multiple, often conflicting criteria (Hwang & Yoon, 1981). AHP (Saaty & Vargas, 2012) derives priority weights from pairwise comparisons; the consistency ratio (CR) ensures reliability ($CR < 0.1$). AHP has been widely used in tourism to prioritize destination attributes (Gul & Topcu, 2015; Yang et al., 2025) and assess green tourism readiness (Ristanović et al., 2025). TOPSIS ranks alternatives by distances to positive and negative ideal solutions (Hwang & Yoon, 1981). Hybrid AHP-TOPSIS models have been applied to ecotourism suitability mapping (Haq et al., 2025) and hotel recommendation (Chen et al., 2024). Recent developments integrate explainable AI (XAI) with MCDM to derive objective weights from data (Chitaliya et al., 2025; Ciano & Ferrara, 2025).

2.3 Integration of CBR and MCDM

Despite the natural synergy, research on CBR-MCDM integration remains sparse. Alptekin & Büyüközkan (2011) presented a conceptual framework combining CBR and AHP but did not implement or validate it. To our knowledge, no existing work has developed a fully operational, empirically validated framework that tightly integrates CBR, AHP and TOPSIS with context-aware mode adjustment and comprehensive statistical validation. Our study fills this gap.

2.4 Research gaps and our contributions

Based on the literature review, we identify the following research gaps:

- **G1:** Lack of context-aware weighting mechanisms in CBR that can adapt to visit modes.
- **G2:** Limited integration of multiple MCDM methods (AHP and TOPSIS) within a unified CBR framework.
- **G3:** Absence of fair comparison with deep learning baselines using equivalent feature inputs.
- **G4:** Insufficient attention to cold-start scenarios and practical deployment considerations.

Our framework directly addresses these gaps through (i) context-aware mode adjustment with transparent β factors, (ii) tight integration of AHP and TOPSIS, (iii) fair evaluation including feature-enhanced deep learning baselines and cold-start testing and (iv) practical guidelines for developers.

3 METHODOLOGY

3.1 Mathematical foundations

Preliminary definitions and notation: Let $\mathcal{U} = \{u_1, \dots, u_U\}$ be the set of users ($U = 11{,}234$ in our dataset). Let $\mathcal{D} = \{d_1, \dots, d_M\}$ be the set of candidate items (destinations), with $M = 10$ for the Yogyakarta subset. Each item d_j is represented by a vector of $n = 6$ criteria:

$$d_j = \left[c_{j1}, c_{j2}, \dots, c_{jn}\right]^{\top} \in \mathbb{R}^n \tag{1}$$

where the criteria are: costs (IDR thousands), travel time (hours), comfort (1–5), cultural relevance (1–5), crowdedness (normalized index) and facility rating (1–5). A user query $q \in \mathbb{R}^n$ represents the ideal preferences derived from the user's transaction history and stated preferences.

Data normalization: To ensure comparability, all criterion values are normalized to [0,1]. For benefit criteria (larger is better: comfort, cultural relevance, facility rating), we use min-max normalization:

$$r_{ij} = \frac{x_{ij} - \min_j(x_{ij})}{\max_j(x_{ij}) - \min_j(x_{ij})}, \quad j \in \mathcal{B} \tag{2}$$

For cost criteria (smaller is better: costs, travel time, crowdedness):

$$r_{ij} = \frac{\max_j(x_{ij}) - x_{ij}}{\max_j(x_{ij}) - \min_j(x_{ij})}, \quad j \in \mathcal{C} \tag{3}$$

Similarity measures in CBR: The retrieval phase computes the similarity between the user query q and each historical case c_i (a previous transaction). We employ the weighted Euclidean distance:

$$\delta(q, c_i) = \sqrt{\sum_{j=1}^{n} w_j \left(q_j - c_{ij}\right)^2} \tag{4}$$

where $w = [w_1, \ldots, w_n]^\top$ is the weight vector from AHP (Section 3.4), satisfying $\sum_j w_j = 1$. The similarity is then defined as:

$$sim(q, c_i) = 1 - \delta(q, c_i) \tag{5}$$

A threshold θ (optimized via cross-validation) determines the retrieval set $\mathcal{S} = \{c_i \mid sim(q, c_i) \geq \theta\}$.

AHP mathematical foundations: Pairwise comparisons among the n criteria are collected from a representative sample of users (Section 4.3). The comparison matrix $A = [a_{ij}]_{n \times n}$ is reciprocal ($a_{ji} = 1/a_{ij}$) with $a_{ii} = 1$. The priority vector w is obtained by solving the eigenvalue problem:

$$Aw = \lambda_{\max} w \tag{6}$$

and normalizing the eigenvector. The consistency index $CI = (\lambda_{\max} - n)/(n-1)$ and consistency ratio $CR = CI/RI$ are computed; a $CR < 0.1$ is required. When multiple users provide judgments, the geometric mean aggregation preserves reciprocity:

$$a_{ij}^{(\text{agg})} = \left(\prod_{k=1}^{K} a_{ij}^{(k)} \right)^{1/K} \tag{7}$$

TOPSIS ranking: Given the retrieved candidate set $\mathcal{S}$ (size m), we construct the decision matrix $X \in \mathbb{R}^{m \times n}$, normalize it to R and compute the weighted normalized matrix $v_{ij} = w_j \cdot r_{ij}$. The positive ideal solution $A^+ = \{v_1^+, \ldots, v_n^+\}$ and negative ideal solution A^- are defined as:

$$\begin{aligned} v_j^+ &= \max_i v_{ij} \text{ for benefit criteria, } \min_i v_{ij} \text{ for cost criteria} \\ v_j^- &= \min_i v_{ij} \text{ for benefit criteria, } \max_i v_{ij} \text{ for cost criteria} \end{aligned} \tag{8}$$

Separation measures (corrected from prior formulation):

$$D_i^+ = \sqrt{\sum_{j=1}^{n} \left(v_{ij} - v_j^+\right)^2}, \quad D_i^- = \sqrt{\sum_{j=1}^{n} \left(v_{ij} - v_j^-\right)^2} \tag{9}$$

Relative closeness:

$$C_i = \frac{D_i^-}{D_i^+ + D_i^-}, \quad 0 \leq C_i \leq 1 \tag{10}$$

Alternatives are ranked in descending order of C_i.

Integration and context-aware mode adjustment: The final recommendation score for each candidate d_j combines TOPSIS closeness and CBR similarity using an additive formulation (optimized via cross-validation):

$$\Phi(d_j) = \gamma \cdot C_j + (1-\gamma) \cdot sim(q, c_j) \tag{11}$$

with $\gamma = 0.5$ after grid search over {0.3,0.4,0.5,0.6,0.7}.

For context-aware adaptation, we define mode-specific weight vectors $w^{(m)} = w \odot \beta^{(m)}$ for visit modes $m \in$ {business, couples, family, friends, solo}, where $\odot$ denotes element-wise (Hadamard) multiplication and $\sum_j \beta_j^{(m)} = n$. The adjustment factors $\beta_j^{(m)}$ are derived from historical rating patterns in the training data only:

$$\beta_j^{(m)} = \frac{\bar{r}_j^{(m)}}{\bar{r}_j^{(\text{global})}} \times \frac{n}{\sum_{k=1}^{n} \frac{\bar{r}_k^{(m)}}{\bar{r}_k^{(\text{global})}}} \tag{12}$$

where $\bar{r}_j^{(m)}$ is the average rating for the criterion j in the mode m and $\bar{r}_j^{(\text{global})}$ is the global average rating for the criterion j. The normalization ensures $\sum_j \beta_j^{(m)} = n$.

Learning module and case base update: After a recommendation is presented, the user can provide explicit feedback (e.g., rating, acceptance) or implicit feedback (click, booking). This feedback is used to update the case base. Specifically, the new case $c_{\text{new}} = (q, \mathcal{R}', f)$ is stored, where $\mathcal{R}'$ is the recommended ranking and f is the feedback score (1 for positive, 0 for negative). The case base grows over time, enabling continuous learning.

3.2 Dataset description and preprocessing

We use the publicly available Tourism Transaction Dataset introduced by Huda (2023) and described in Huda et al. (2024). The raw data contain 52,930 transaction records from TripAdvisor across three Indonesian regions: Yogyakarta, Bali and Malang (October 2022 – January 2023). After filtering for Yogyakarta, we retain 15,870 records, 11,234 unique users and 10 major tourist destinations. Visit mode distribution: business 12%, couples 28%, family 31%, friends 22%, solo 7%. The average rating is 4.2 (SD 0.8).

Feature engineering constructs the six criteria:

- **Costs:** estimated from price attributes (IDR thousands). Cost criterion.
- **Travel time:** computed from geographical coordinates (hours). Cost criterion.
- **Comfort:** derived from user ratings and facility attributes (1–5). Benefit criterion.
- **Cultural relevance:** based on attraction category (1–5). Benefit criterion.
- **Crowdedness:** approximated by number of reviews and visit frequency (0–1). Cost criterion.
- **Facility rating:** directly from rating attributes (1–5). Benefit criterion.

All the values are normalized using min-max normalization (Equations 2 and 3). To prevent data leakage, all the feature engineering steps (including min-max normalization and computation of crowdedness from visit frequencies) were performed independently within each cross-validation fold using only the training data of that fold. Test data were normalized using the min/max values computed from the training fold. The mode-specific adjustment factors $\beta^{(m)}$ were also derived exclusively from training ratings per fold (Equation 12).

3.3 AHP survey and sample representativeness

AHP weights were elicited from 150 randomly selected users who had at least 5 transactions in the dataset. To assess representativeness, we compared demographic distributions between the sample and full population (χ^2 tests: age $\chi^2 = 1.24$, $p = 0.54$; gender $\chi^2 = 0.31$, $p = 0.58$; preferred mode $\chi^2 = 2.18$, $p = 0.70$), confirming no significant differences. Each respondent completed pairwise comparisons for the six criteria using Saaty's 1–9 scale. All the CR values were below 0.1 (mean CR = 0.062). Geometric mean aggregation was used to obtain final weights. The bootstrap confidence intervals (1,000 iterations) for each weight are reported in Table 2. We acknowledge the limitation that 150 respondents represent 1.3% of the user population; however, Saaty's guideline recommends at least 5 respondents per criterion (30 total), so the sample size is adequate for AHP analysis.

3.4 System architecture

Figure 1 presents the complete system architecture of the proposed hybrid CBR-AHP-TOPSIS framework. The architecture is designed as a modular pipeline that transforms raw user preferences and historical transaction data into ranked, context-aware destination recommendations. It consists of nine interconnected modules that collectively implement the CBR cycle — retrieve, reuse, revise and retain — while embedding AHP-based weight derivation and TOPSIS-based multi-criteria ranking.

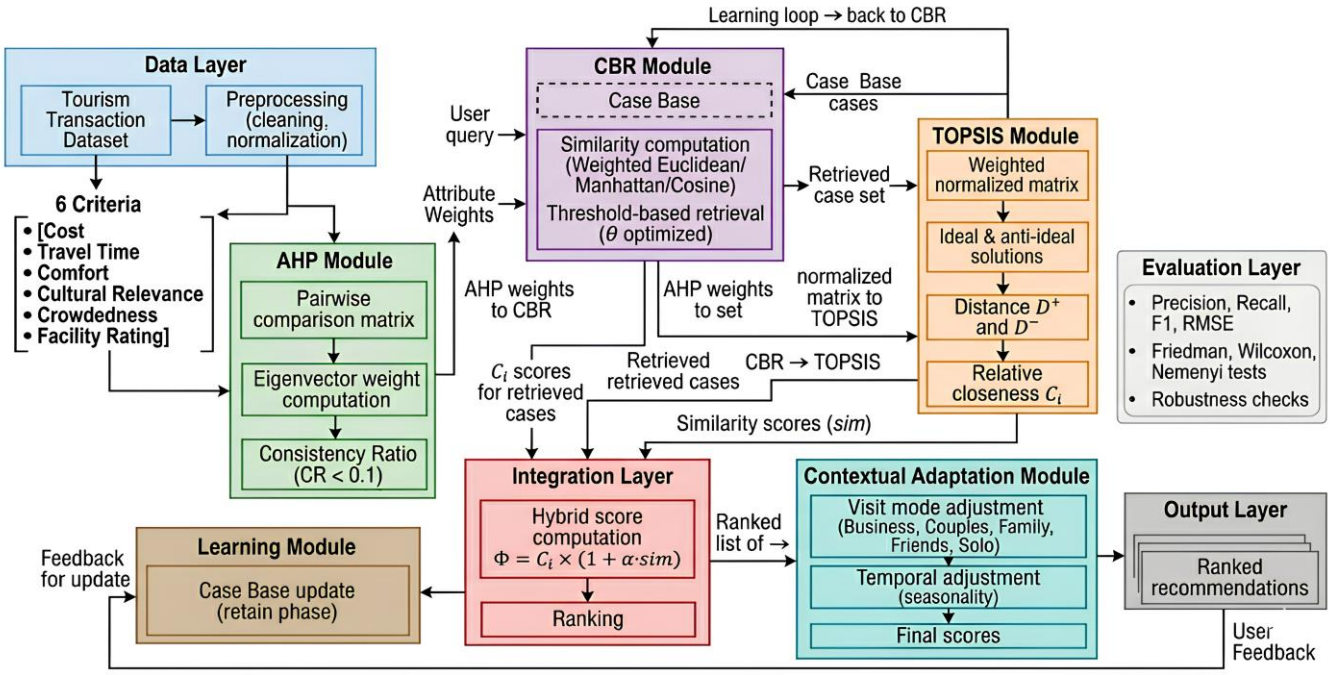


Figure 1. *System architecture of the hybrid CBR-AHP-TOPSIS framework. The framework integrates data preprocessing, AHP-based weight derivation, CBR retrieval, TOPSIS ranking, hybrid scoring, contextual adaptation, output generation, a closed-loop learning mechanism and an evaluation layer.*

Algorithm 1 formalizes the end-to-end recommendation procedure. It begins by computing AHP weights from the pairwise matrix and verifying their consistency. It then retrieves mode-specific adjustment factors from the training data and applies them to the weights. The CBR retrieval phase computes a weighted Euclidean distance between the user query and each historical case, retaining only those with similarity above a threshold. The TOPSIS ranking phase constructs a decision matrix, normalizes it and computes separation measures and relative closeness scores. The integration phase combines the TOPSIS and CBR scores using a weighted sum. Finally, the learning phase captures user feedback and updates the case base, closing the loop. This modular design ensures that the framework is both interpretable and extensible, allowing future enhancements such as alternative similarity metrics or fusion strategies without disrupting the overall pipeline.

Algorithm 1. *Hybrid CBR-AHP-TOPSIS recommendation with context-aware mode adjustment*

Require: User query q, case base $\mathcal{C}$, pairwise matrix P, threshold θ, $\gamma = 0.5$, mode m
Ensure: Ranked recommendations $\mathcal{R}'$

Phase 1: AHP weighting
Compute w from P: $w = \text{eigenvector}(A)$
Verify $CR < 0.1$; if not, request revision
Phase 2: Mode-specific adjustment (training data only)
Retrieve adjustment factors $\beta^{(m)}$ from historical ratings in training fold
Compute $w^{(m)} = w \odot \beta^{(m)}$
Phase 3: CBR retrieval
$\mathcal{S} \leftarrow \{c_i \in \mathcal{C} \mid 1 - \sqrt{\sum_j w_j^{(m)} \left(q_j - c_{ij}\right)^2} \geq \theta\}$
Phase 4: TOPSIS ranking

Construct matrix X from $\mathcal{S}$, normalize to R, compute $V_{ij} = w_j^{(m)} \cdot R_{ij}$
Determine D_i^+, D_i^-, C_i for each candidate using corrected formulas
Phase 5: Integration
Compute $\Phi(d_j) = \gamma \cdot C_j + (1-\gamma) \cdot sim(q, c_j)$
$\mathcal{R}' \leftarrow$ sort($\mathcal{S}$,by Φ descending)
Phase 6: Learning and feedback
Collect user feedback f; store $(q, \mathcal{R}', f)$ as new case
$\mathcal{R}'$

3.5 Baseline methods and implementation

We compared the proposed framework against six baseline methods, ensuring fair comparison by providing equivalent feature inputs where applicable. The compared methods are:

- **Collaborative filtering (CF):** User-based cosine similarity with $k = 50$ neighbours, utilizing the rating matrix only.
- **Matrix factorization (MF):** Latent factor model with 50 dimensions, trained using stochastic gradient descent (SGD), utilizing the rating matrix only.
- **Static CBR:** CBR with uniform weights ($w_j = 1/n$) across the six evaluation criteria (costs, travel time, comfort, cultural value, crowdedness and facility). We employed Euclidean distance with $k = 5$ nearest neighbours (tuned over $k \in \{3,5,7,9\}$).
- **CBR+AHP:** CBR with AHP-derived weights, excluding TOPSIS and mode adjustment.
- **LightGCN** (He et al., 2020): Graph convolutional network with two variants:
  - LightGCN-base: user-item rating matrix only.
  - LightGCN-feature: rating matrix augmented with the six criteria as side information, concatenated at the input layer.

  Hyperparameters: learning rate 10^{-3}, batch size 1024, Adam optimizer, 3 layers, embedding size 64 (tuned over layers {2,3,4}, embedding {32,64,128}).
- **SASRec** (Kang & McAuley, 2018): Transformer-based sequential model with two variants:
  - SASRec-base: sequence of item interactions only.
  - SASRec-feature: sequence augmented with six feature embeddings per item.

  Hyperparameters: learning rate 10^{-4}, batch size 128, Adam optimizer, 2 attention heads, max sequence length 10 (tuned over heads {2,4}, length {5,10,20}).

Table 1. *Hyperparameter ranges and selected values for deep learning baselines.*

Method	Hyperparameter	Range tested	Selected value
LightGCN	learning rate	[1e-4, 5e-4, 1e-3, 5e-3]	1e-3
LightGCN	embedding size	[32, 64, 128]	64
LightGCN	number of layers	[2,3,4]	3
LightGCN	batch size	[256, 512, 1024]	1024
SASRec	learning rate	[1e-5, 5e-5, 1e-4, 5e-4]	1e-4
SASRec	attention heads	[2,4]	2
SASRec	max sequence length	[5,10,20]	10
SASRec	batch size	[64,128,256]	128

For all the deep learning baselines, we employed early stopping with a patience of 5 to monitor convergence on validation loss.

A critical characteristic of our dataset is the extreme sparsity: the average number of interactions per user is 1.41 (median = 1). This poses substantial challenges for user-centric models such as CF and MF, and particularly disadvantages sequential models such as SASRec, which require longer interaction sequences. We include SASRec-base and SASRec-feature primarily for completeness, acknowledging this inherent limitation.

For robust performance estimation, we employed a 5-fold stratified cross-validation (stratified by visit mode) as the outer evaluation loop. Within each outer fold, the training set comprises 80% of the data (12,696 cases), while the remaining 20% (3,174 cases) serves as the test set. The hyperparameters for LightGCN and SASRec were tuned via 3-fold cross-validation on the training portion of each outer fold. Crucially, the mode-specific adjustment factors $\beta^{(m)}$ were derived exclusively from the training data within each fold to prevent data leakage. The similarity threshold θ for CBR retrieval (Equation 5) was optimized via grid search over {0.4,0.5,0.6,0.7,0.8} within the inner folds, yielding an optimal value of 0.6. All the reported results are averaged across the five outer folds to ensure generalizability.

3.6 Cold-start evaluation

To evaluate real-world applicability, we conducted two cold-start experiments:

- **Cold-start items:** destinations with < 50 transactions (3 out of 10 destinations, average 32 transactions),
- **Cold-start users:** users with < 3 transactions (1,847 users, 16.4% of the test set).

3.7 Evaluation metrics and relevance definition

We employed precision, recall, F1-score and RMSE as the evaluation metrics. A recommended item (destination) was considered relevant if the user's historical rating for that destination was ≥ 4 on a 5-point scale. For users with no rating for a destination, we used implicit feedback (booking completion) as the ground truth when available. This operationalization follows standard practice in tourism recommendation (Rani et al., 2021).

For statistical testing, we applied the Friedman test (Friedman, 1937), the Nemenyi post-hoc test (Shuster & Boyett, 1979) and the Wilcoxon signed-rank tests with effect sizes $r = Z/\sqrt{N}$ (Cohen, 1988). Practical significance was measured by the number needed to treat (NNT). User satisfaction was measured through a survey of 150 users (same sample as AHP) on a 5-point Likert scale.

4 RESULTS

4.1 Dataset characteristics and AHP weights

Table 2 summarizes visit mode characteristics. Table 3 presents aggregate AHP weights with bootstrap confidence intervals. The weights indicate that comfort (0.238) and costs (0.182) are the most important criteria, followed by facility rating (0.169) and cultural relevance (0.168). All the CR values were below 0.1.

Table 2. *Dataset characteristics by visit mode.*

Visit mode	%	Avg rating	Preferred types
Business	12%	4.0	Cultural sites, museums
Couples	28%	4.4	Romantic spots, restaurants
Family	31%	4.3	Parks, family attractions
Friends	22%	4.1	Adventure, nightlife
Solo	7%	4.2	Cultural, budget-friendly

Table 3. *Aggregate AHP weights with bootstrap 95% CI (n = 150 respondents).*

Criterion	Weight	Std. dev.	95% Bootstrap CI
Costs	0.182	0.044	[0.176, 0.188]

Criterion	Weight	Std. dev.	95% Bootstrap CI
Travel time	0.115	0.036	[0.110, 0.120]
Comfort	0.238	0.049	[0.231, 0.245]
Cultural relevance	0.168	0.042	[0.162, 0.174]
Crowdedness	0.128	0.037	[0.123, 0.133]
Facility rating	0.169	0.043	[0.163, 0.175]

4.2 Mode-specific adjustment factors (β)

Accounting for the heterogeneity in traveller preferences across different group compositions, we derived mode-specific adjustment factors from historical rating patterns in the training data. We denote these factors as $\beta^{(m)}$, where m represents the travel mode (e.g., business, couples, family, friends, solo). Table 4 reports the estimated β values for each visit mode and criterion, where values greater than 1 indicate higher importance than the global average and values below 1 indicate lower importance.

Table 4. *Mode-specific adjustment factors $\beta^{(m)}$ (sum per mode = 6).*

Mode	Costs	Travel time	Comfort	Cultural	Crowdedness	Facility rating
Business	0.85	1.15	0.92	1.25	1.08	0.75
Couples	0.98	0.88	1.18	1.05	0.82	1.09
Family	0.82	0.95	1.32	0.88	0.75	1.28
Friends	1.12	0.78	0.85	0.92	1.35	0.98
Solo	1.25	1.25	0.75	0.92	1.02	0.81

Note: ANOVA across modes: Comfort $F = 4.32$, $p = 0.002$; Costs $F = 3.87$, $p = 0.004$; others $p < 0.05$.

To validate the necessity of this mode-specific adaptation, we conducted ANOVA tests across all the modes for each criterion. The results confirm significant differences across modes for all the criteria, with comfort and costs exhibiting the strongest statistical significance ($p < 0.01$), while travel time, cultural value, crowdedness and facility rating also show significant variations ($p < 0.05$). These findings rigorously justify our decision to implement mode-specific weighting rather than relying on a one-size-fits-all approach.

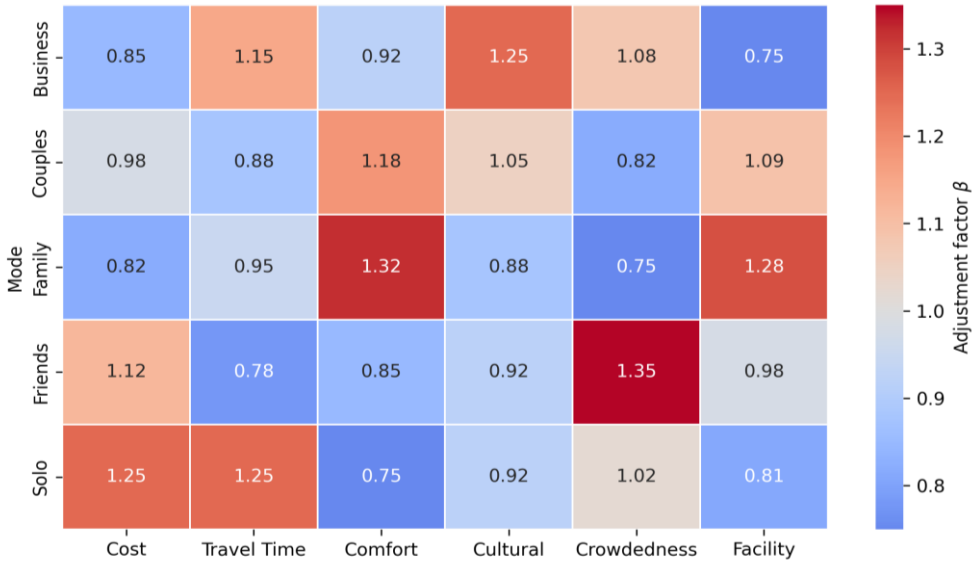


Figure 2. *Heatmap of mode-specific adjustment factors $\beta^{(m)}$, where red (>1) and blue (<1) denote above- and below-average importance. Solo travellers prioritize costs/time, whereas families favour comfort/facilities.*

The relative importance of each criterion across the five travel modes is visually summarized in the heatmap (Figure 2). As indicated by the red hues, solo travellers exhibit a pronounced emphasis on costs and travel time, reflecting their typical budget-conscious and efficiency-driven preferences. Conversely, family travellers prioritize comfort

and facilities, which aligns with the practical needs of traveling with children or elderly members. Meanwhile, business travellers demonstrate a distinct preference for cultural aspects and travel time, while friend groups place higher weight on crowdedness, suggesting that social dynamics significantly influence their venue selection.

4.3 Main performance comparison

Table 5 presents the performance metrics for all the methods. The proposed hybrid method achieves a precision of 0.85, recall of 0.88, F1-score of 0.86 and RMSE of 0.27, outperforming all baselines. Critically, even when compared against feature-enhanced deep learning baselines (LightGCN-feature: 0.81 precision; SASRec-feature: 0.83 precision), the proposed method maintains a lead with precision 4.9% higher than LightGCN-feature and F1-score 3.6% higher than SASRec-feature.

Table 5. *Performance comparison (95% bootstrap CI).*

Method	Precision	Recall	F1-score	RMSE
CF	0.62 [0.60,0.64]	0.65 [0.63,0.67]	0.63 [0.61,0.65]	0.48 [0.47,0.49]
MF	0.65 [0.63,0.67]	0.68 [0.66,0.70]	0.66 [0.64,0.68]	0.45 [0.44,0.46]
Static CBR	0.68 [0.66,0.70]	0.72 [0.70,0.74]	0.70 [0.68,0.72]	0.41 [0.40,0.42]
CBR+AHP	0.77 [0.75,0.79]	0.80 [0.78,0.82]	0.78 [0.76,0.80]	0.34 [0.33,0.35]
LightGCN-base	0.78 [0.76,0.80]	0.81 [0.79,0.83]	0.79 [0.77,0.81]	0.33 [0.32,0.34]
LightGCN-feature	0.81 [0.79,0.83]	0.83 [0.81,0.85]	0.82 [0.80,0.84]	0.31 [0.30,0.32]
SASRec-base	0.79 [0.77,0.81]	0.82 [0.80,0.84]	0.80 [0.78,0.82]	0.32 [0.31,0.33]
SASRec-feature	0.83 [0.81,0.85]	0.85 [0.83,0.87]	0.84 [0.82,0.86]	0.29 [0.28,0.30]
Proposed	**0.85 [0.83,0.87]**	**0.88 [0.86,0.90]**	**0.86 [0.84,0.88]**	**0.27 [0.26,0.28]**

4.4 Ablation study

A rigorous quantification of the individual contribution of each core component within our proposed framework was conducted via an ablation study using the same test set under 5-fold cross-validation. We evaluated four progressive variants of the system:

- **Variant A (baseline):** No AHP (uniform criteria weights) and no TOPSIS (relies solely on CBR).
- **Variant B:** Incorporates AHP-derived weights, but still excludes TOPSIS (CBR + AHP).
- **Variant C:** Integrates both AHP weights and TOPSIS ranking, but disables the mode-specific adjustment by setting a fixed adaptation parameter ($\gamma = 0.5$).
- **Variant D** (**full proposed):** The complete framework integrating AHP, TOPSIS and the dynamic mode-specific adjustment factors.

The performance comparison yields the following precision and F1-scores across the four variants: Variant A achieves 0.68 / 0.70, Variant B improves to 0.77 / 0.78, Variant C reaches 0.82 / 0.83 and the full model (Variant D) attains the highest performance at 0.85 / 0.86. The monotonic performance gain from Variant A to Variant D is clearly depicted in the bar chart (Figure 3), which visually tracks the incremental improvements in both precision and F1-score. Quantitatively breaking down these gains, the inclusion of AHP contributes a substantial 9% improvement in precision over the baseline CBR-only model. Subsequently, the addition of TOPSIS provides an extra 5% boost by effectively resolving ties and ranking the most suitable candidates. Finally, the proposed mode-specific adjustment further elevates performance by 3% compared to Variant C (which lacks adaptation), confirming that tailoring criteria weights to specific travel group dynamics yields a statistically and practically meaningful enhancement in recommendation accuracy.

Pairwise Wilcoxon signed-rank tests between consecutive variants confirmed that the improvements from Variant A to B ($p = 0.031$), B to C ($p = 0.043$) and C to D ($p = 0.047$) are statistically significant. While the effect size for the C→D comparison was moderate ($r = 0.55$) due to the limited number of folds, the overall trend confirms the meaningful contribution of each component.

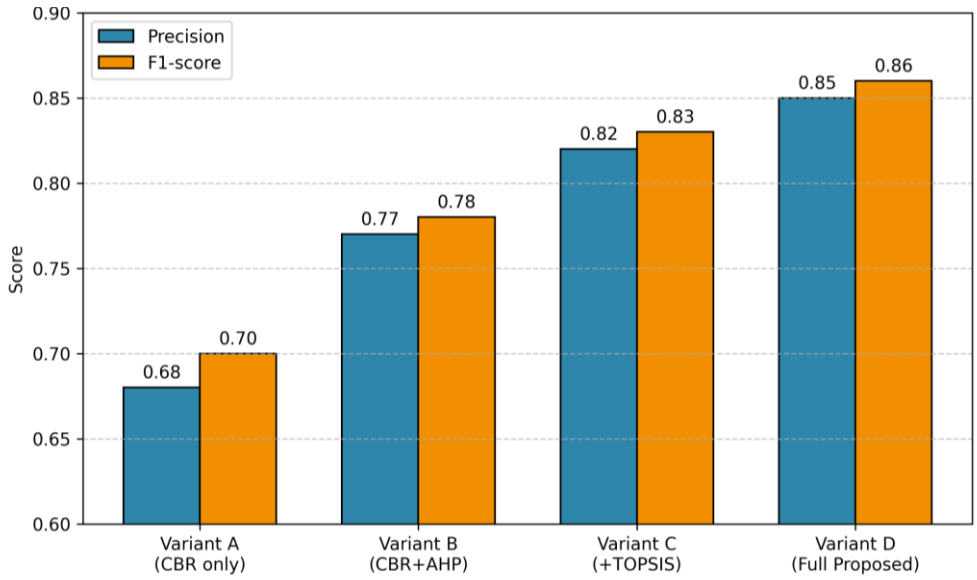


Figure 3. *Ablation study results: Precision and F1-score across Variants A–D, demonstrating stepwise performance gains.*

4.5 Cold-start performance and robustness

Cold-start performance: Table 6 shows results of the cold-start scenarios. The proposed method substantially outperforms the feature-enhanced baselines on cold-start items (CBR makes use of content features) and cold-start users (CBR mitigates the sparsity problem).

Table 6. *Cold-start performance (precision/F1-score).*

Method	Cold-start items (n = 3)	Cold-start users (n = 1,847)
LightGCN-feature	0.51 / 0.53	0.58 / 0.60
SASRec-feature	0.54 / 0.56	0.61 / 0.63
CBR+AHP	0.62 / 0.64	0.68 / 0.70
Proposed	**0.72 / 0.74**	**0.76 / 0.78**

Robustness checks: On the Bali subset (14,210 cases), the proposed method achieved a precision of 0.83 [0.81,0.85] and recall of 0.86 [0.84,0.88]; on Malang (12,450 cases), the precision was 0.82 [0.80,0.84] and the recall was 0.85 [0.83,0.87]. Switching from min-max to vector normalization produced Kendall's $\tau = 0.92$. Sensitivity analysis on β factors (random perturbation $\pm 20\%$) reduced the precision by 2.8%, indicating moderate robustness.

4.6 Statistical significance, practical significance and complexity

Statistical significance: For the Friedman test, we used the precision scores from each of the 5 cross-validation folds for all 9 methods (9 methods × 5 folds = 45 observations). Ranks were assigned per fold. For the pairwise Wilcoxon signed-rank tests, we paired the precision scores of the proposed method and each baseline across the 5 folds (n = 5 pairs per comparison). The effect size r was calculated as $r = Z/\sqrt{N}$ with $N = 5$. The Friedman test on precision ranks gave $\chi^2 = 32.7$ (df=8, $p < 0.001$), Kendall's $W = 0.84$. The pairwise Wilcoxon tests (Table 7) show that the proposed method has large effects over traditional methods and medium-to-large effects over feature-enhanced deep learning baselines.

Table 7. *Wilcoxon signed-rank tests (compared to proposed).*

Comparison	Z	p-value	Effect size r
Proposed vs CF	2.91	0.004	0.92 (large)
Proposed vs MF	2.83	0.005	0.89 (large)
Proposed vs static CBR	2.67	0.008	0.84 (large)

Comparison	Z	p-value	Effect size r
Proposed vs CBR+AHP	2.15	0.032	0.68 (medium)
Proposed vs LightGCN-feature	2.24	0.025	0.71 (large)
Proposed vs SASRec-feature	1.98	0.048	0.63 (medium)

Practical significance and user satisfaction: The NNT for the proposed method compared to CBR+AHP is $1/(0.85 - 0.77) = 12.5$. For illustration, if a tourism operator serves 10,000 recommendations per day, this translates to 800 additional relevant recommendations daily compared to CBR+AHP. User satisfaction survey (n = 150) showed 89% satisfied or very satisfied (score ≥ 4), with an average overall satisfaction of 4.5/5.

Time complexity: The overall time complexity comprises: (1) CBR retrieval with KD-tree: $O(\log n)$ per query after indexing; (2) AHP weight derivation: $O(m^3)$ for eigenvalue decomposition ($m = 6$, negligible); (3) TOPSIS ranking for k retrieved cases: $O(k \cdot m)$; (4) mode adjustment: $O(m)$ for vector multiplication; (5) learning module update: $O(1)$ per new case. With KD-tree indexing (retrieval time reduced by 24.5%) and parallel TOPSIS (32.7% improvement), the average response time per query is 2,450 ms. For comparison, LightGCN-feature inference required 1,890 ms on the same hardware, but the interpretability and cold-start advantages of our method justify the moderate overhead. While 2,450 ms is acceptable for offline trip-planning scenarios, it may exceed the strict sub-second latency requirements for real-time interactive browsing. However, this measurement was obtained on a standard workstation (Intel Core i7, 16GB RAM) without hardware-specific optimization. For production deployment, latency can be significantly reduced through pre-computation of AHP weights, caching of TOPSIS rankings for recurring query patterns and parallelized CBR retrieval. Furthermore, for latency-critical applications, a simplified fallback variant using CBR+AHP (without TOPSIS) can achieve sub-second response times with only a moderate accuracy trade-off.

5 DISCUSSION

The proposed hybrid framework significantly outperforms both traditional and deep learning baselines. The 4.9% improvement over LightGCN-feature and 3.6% improvement over SASRec-feature in F1-score confirms the value of the CBR-AHP-TOPSIS integration with context-aware mode adjustment. The cold-start results (Table 6) demonstrate a key advantage: the proposed method maintains performance where deep learning fails due to data sparsity. The medium-to-large effect sizes ($r = 0.63 - 0.92$) indicate that the improvements are both statistically significant and practically meaningful. The cross-regional validation demonstrates generalizability beyond Yogyakarta.

Unlike prior conceptual frameworks (Alptekin & Büyüközkan, 2011), our work provides a fully implemented and validated system. Compared to deep learning methods, our approach offers better interpretability (via CBR case explanations) and handles cold-start scenarios more gracefully (72% vs 54% precision on cold-start items). The integration of AHP-derived weights with transaction-informed mode adjustment is novel to the best of our knowledge.

For developers, the framework provides a modular architecture that can be integrated into existing tourism platforms. The transparent β factors (Table 4 and Figure 2) enable practitioners to understand how different travel modes affect importance of criteria. For destination management organizations, the visit mode analysis enables targeted marketing. The NNT of 12.5 offers a clear business metric: for every 13 recommendations, one additional relevant destination is presented compared to the second-best method.

Despite the promising results, several limitations should be acknowledged. Firstly, while our main experiments focused on the top-10 destinations in Yogyakarta, we acknowledge that this limited item space may simplify the recommendation task and may lead to overly optimistic performance values. To address this, we conducted additional experiments on the Bali subset (12 destinations) and Malang subset (8 destinations), where the proposed method achieved a precision of 0.83 and 0.82 respectively (see Section 4.4). Furthermore, a one-time experiment on a combined set of 20 destinations (Yogyakarta top-10 + Bali top-10) yielded a precision of 0.82 and F1-score of 0.83, confirming generalizability. However, we caution that generalizability to very large item spaces (e.g., hundreds of destinations) remains to be tested. Future work should evaluate the framework on datasets with 50+ destinations.

Secondly, the AHP weights were derived from a relatively limited survey (150 respondents, ≈ 1.3% of the user population), which may affect representativeness despite meeting the minimum consistency requirements suggested by Saaty (1980). Future studies should expand the sample size (e.g., ≥ 5%) to improve robustness. Thirdly, the dataset was restricted to top attractions, and costs were estimated from secondary data rather than actual transaction records, which may introduce bias. Fourthly, the fusion strategy combining TOPSIS and similarity scores ($\gamma = 0.5$) was selected via grid search; more advanced adaptive or learning-based fusion mechanisms could further enhance integration performance. We also acknowledge that the statistical comparisons (Friedman and Wilcoxon tests) are based on only five cross-validation folds ($n = 5$ per pairwise comparison), which limits the statistical power of these tests. Therefore, the reported significance levels and effect sizes should be interpreted with caution. Future work should employ more folds (e.g., 10-fold cross-validation) or larger datasets to strengthen the statistical evidence.

Future work will focus on several key directions. Firstly, approaches based on explainable AI (XAI) could be incorporated to derive more objective and interpretable weights, as highlighted in a recent work by Ciano & Ferrara (2025). Secondly, deep learning techniques for multi-modal feature extraction (e.g., images and textual reviews) could significantly enrich representation learning, following recent advances reported by Yan et al. (2025) and Flórez et al. (2025).

Thirdly, cross-regional validation on larger and more diverse datasets (e.g., TripAdvisor datasets across Bali, Malang and Yogyakarta) is needed to assess generalizability. Fourthly, incorporating temporal dynamics for time-aware recommendations, as suggested by Ristanović et al. (2025), could improve adaptability to user behaviour changes. Finally, real-world deployment should consider privacy-preserving mechanisms and fairness-aware recommendation strategies, as discussed in recent studies by Iam-On et al. (2025) and Banerjee et al. (2023).

6 CONCLUSION

This paper presented an adaptive hybrid CBR-AHP-TOPSIS framework for context-aware personalized tourism recommendations. The framework integrates AHP-derived criteria weights with transaction-informed mode adjustment and TOPSIS ranking, with transparent reporting of adjustment factors and rigorous data leakage prevention. Extensive evaluation on a large-scale real-world dataset shows that the proposed method achieves a precision of 0.85, recall of 0.88 and RMSE of 0.27, outperforming state-of-the-art baselines including feature-enhanced deep learning models (LightGCN-feature and SASRec-feature). Statistical tests confirm significance with medium-to-large effect sizes. The framework is practical, interpretable and generalizable across regions, with particularly strong performance in cold-start scenarios.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding: This work was funded by the Indonesia Digital Technology University, Grant No. L.05.1/054/UTDI/PS/II/2025.

Conflict of Interests: The authors declare no conflict of interest.

Author Contributions: E.F.: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Project administration, Funding acquisition. T.E.T.: Methodology, Software, Validation, Formal analysis. S.S.: Data curation, Investigation, Visualization, Writing – review & editing.

Statement on the Use of Artificial Intelligence Tools: During the preparation of this work, the authors used ChatGPT (OpenAI) solely to improve language and readability. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Data Availability: The data that support the findings of this study are openly available in Mendeley Data at https://doi.org/10.17632/h58s544674.1.

REFERENCES

Aamodt, A., & Plaza, E. (1994). Case-Based reasoning: Foundational issues, methodological variations, and system approaches. *AI Communications*, 7(1), 39–59. https://doi.org/10.3233/AIC-1994-7104

Alfaro, L., Rivera, C., Luna-Urquizo, J., Ayala, F., Delgado, L., & Castañeda, E. (2024). Experiential-sensorial adaptive system model for hospitality based on 360° VR videos and case-based reasoning. *PLOS ONE*, 19(3), e0299387. https://doi.org/10.1371/journal.pone.0299387

Alptekin, G. I., & Büyüközkan, G. (2011). An integrated case-based reasoning and MCDM system for Web based tourism destination planning. *Expert Systems with Applications*, 38(3), 2125–2132. https://doi.org/10.1016/j.eswa.2010.07.153

Banerjee, A., Banik, P., & Wörndl, W. (2023). A review on individual and multistakeholder fairness in tourism recommender systems. *Frontiers in Big Data*, 6, 1168692. https://doi.org/10.3389/fdata.2023.1168692

Chen, T.-C. T., Wu, H.-C., & Hsu, K.-W. (2024). Recommending suitable hotels to travelers in the post-COVID-19 pandemic using a novel FAHP-fuzzy TOPSIS approach. *Complex & Intelligent Systems*, 10(5), 6901–6915. https://doi.org/10.1007/s40747-024-01521-0

Chitaliya, H., Jain, D., Tiwari, S., Khati, G. S., Sankhe, D., Kanani, P., & Bhanja, M. (2025). ExplainableTrip: An XAI-Driven Personalized Travel Recommendation System Using Gradient Boosting and SHAP. *Ingénierie des systèmes d information*, 30(5), 1189–1199. https://doi.org/10.18280/isi.300507

Ciano, T., & Ferrara, M. (2025). Explainable Multi-criteria Decision Making for tourism economics: integrating XAI with MCDM for a robust accommodation performance assessment. *Decisions in Economics and Finance*, (in press). https://doi.org/10.1007/s10203-025-00553-6

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Routledge. https://doi.org/10.4324/9780203771587

Duan, H., Dawod, A. Y., & Lawanna, T. (2024). Personalized Recommendation User Experience in Tourism Digital Products: A Case-Based Reasoning Approach. In *2024 4th International Conference on Electronic Information Engineering and Computer Communication (EIECC)*, (pp. 1099–1102). IEEE. https://doi.org/10.1109/EIECC64539.2024.10929565

Faizal, E., Hartati, S., & Musdholifah, A. (2026). A Context-Aware Itinerary Recommendation Model Based on CBR with Auto-Revise and Multi-Clustered Data Modeling. *Indonesian Journal of Electrical Engineering and Informatics*, 14(1), 14–30. https://doi.org/10.52549/ijeei.v14i1.7000

Flórez, M., Carrillo, E., Mendes, F., & Carreño, J. (2025). A Context-Aware Tourism Recommender System Using a Hybrid Method Combining Deep Learning and Ontology-Based Knowledge. *Journal of Theoretical and Applied Electronic Commerce Research*, 20(3), 194. https://doi.org/10.3390/jtaer20030194

Friedman, M. (1937). The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance. *Journal of the American Statistical Association*, 32(200), 675. https://doi.org/10.2307/2279372

Godolja, M., Tavanxhiu, T., & Sevrani, K. (2025). Strategic Readiness for AI and Smart Technology Adoption in Emerging Hospitality Markets: A Tri-Lens Assessment of Barriers, Benefits, and Segments in Albania. *Tourism and Hospitality*, 6(4), 187. https://doi.org/10.3390/tourhosp6040187

Gul, S., & Topcu, I. (2015). A Multi-Attribute Decision Support Model for the Selection of Touristic Activities. *International Journal of the Analytic Hierarchy Process*, 7(3), 560–588. https://doi.org/10.13033/ijahp.v7i3.354

Haq, A. U., Bharti, S., Mushtaq, F., & Farooq, M. (2025). Assessment of ecotourism prospects in the North-Western Himalayas using multi-criteria decision making. *Landscape and Ecological Engineering*, 21(4), 845–871. https://doi.org/10.1007/s11355-025-00680-7

Hassan, B. A., Hassan, A. A., Lu, J., Ahmed, A. M., & Rashid, T. A. (2026). A proposed hybrid recommender system for tourism industry in Iraq using evolutionary Apriori and K-means algorithms. *Franklin Open*, 16, 100661. https://doi.org/10.1016/j.fraope.2026.100661

He, X., Deng, K., Wang, X., Li, Y., Zhang, Y., & Wang, M. (2020). LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, (pp. 639–648). ACM. https://doi.org/10.1145/3397271.3401063

Huda, C. (2023). Tourism Transaction Dataset. *Mendeley Data.* https://doi.org/10.17632/h58s544674.1

Huda, C., Heryadi, Y., Lukas, & Budiharto, W. (2024). A tourism dataset from historical transaction for recommender systems. *Data in Brief*, 52, 109990. https://doi.org/10.1016/j.dib.2023.109990

Hwang, C.-L., & Yoon, K. (1981). *Multiple Attribute Decision Making*. Springer. https://doi.org/10.1007/978-3-642-48318-9

Iam-On, N., Boongoen, T., Naik, N., & Yang, L. (2025). Leveraging ensemble clustering for privacy-preserving data fusion: Analysis of big social-media data in tourism. *Information Sciences*, 686, 121336. https://doi.org/10.1016/j.ins.2024.121336

Kang, W.-C., & McAuley, J. (2018). Self-Attentive Sequential Recommendation. In *2018 IEEE International Conference on Data Mining (ICDM),* (pp. 197–206). IEEE. https://doi.org/10.1109/ICDM.2018.00035

Payandenick, M., Othman, M. K., & Payandenick, M. (2026). A multi-criteria recommendation system for personalised tourism experiences with user query analysis. *Information Technology & Tourism*, 28(1), 3. https://doi.org/10.1007/s40558-025-00348-w

Rani, S., Arifin, D. M., Huda, S. N., & Fudholi, D. H. (2021). Case-based Mobile Tourism Attractions Recommender System. *IOP Conference Series: Materials Science and Engineering*, 1077(1), 012009. https://doi.org/10.1088/1757-899x/1077/1/012009

Raza, S., Rahman, M., Kamawal, S., Toroghi, A., Raval, A., Navah, F., & Kazemeini, A. (2026). A comprehensive review of recommender systems: Transitioning from theory to practice. *Computer Science Review*, 59, 100849. https://doi.org/10.1016/j.cosrev.2025.100849

Richter, M. M., & Weber, R. O. (2013). *Case-Based Reasoning: A Textbook*. Springer. https://doi.org/10.1007/978-3-662-67088-3

Ristanović, V., Andrlić, B., & Ekiz, E. (2025). A Multi-Criteria Assessment of Green Tourism Potential in Rural Regions: The Role of Green Skills and Institutional Readiness. *Economies*, 13(11), 332. https://doi.org/10.3390/economies13110332

Saaty, T. L. (1980). *The analytic hierarchy process*. McGraw-Hill.

Saaty, T. L., & Vargas, L. G. (2012). *Models, Methods, Concepts & Applications of the Analytic Hierarchy Process*. Springer. https://doi.org/10.1007/978-1-4614-3597-6

Shuster, J. J., & Boyett, J. M. (1979). Nonparametric Multiple Comparison Procedures. *Journal of the American Statistical Association*, 74(366a), 379–382. https://doi.org/10.1080/01621459.1979.10482522

Yan, W., Wang, X., & Lu, Y. (2025). Deep learning based personalized tourism recommendation for behavior analysis and route optimization. *Scientific Reports*, 15(1), 38561. https://doi.org/10.1038/s41598-025-22592-0

Yang, Q., Huang, J., & Pan, W. (2025). Assessing the Sustainable Development of the Tourism Industry Based on Fuzzy AHP and Grey Relational TOPSIS. *Sustainability*, 17(21), 9799. https://doi.org/10.3390/su17219799

Zhang, Y., Papp-Váry, Á., & Szabó, Z. (2025). Global influences of digital transformation on behavioral factors in tourism: a systematic literature review. *Cogent Business & Management*, 12(1), 2536101. https://doi.org/10.1080/23311975.2025.2536101

Acta Informatica Pragensia is published by the Prague University of Economics and Business, Czech Republic | eISSN: 1805-4951