International Journal of Smart Business and Technology
Vol. 14, No. 1 (2026), pp. 81-96
http://dx.doi.org/10.21742/ijsbt.2026.14.1.06

# An Empirically Grounded Location-Aware Framework for Event-Centric Consumer Behavior Modeling Using Social Media Analytics

Mohd Azhar[1*] and Seema Nagar[2]

[1]*Department of Marketing, New Delhi Institute of Management, New Delhi, India*
[2]*IBM Research, Yorktown Heights, NY, United States*
[1]*mohd.azhar@ndimdelhi.org*, [2]*seema.nagar@ibmresearch.com*

## *Abstract*

*The increasing integration of social media into everyday communication has opened new avenues for examining consumer behavior in dynamic, event-centric environments. In particular, large-scale international events generate substantial volumes of user-generated content, offering opportunities to derive context-aware insights for strategic marketing interventions. In this context, the present study proposes an analytics-driven, location-aware consumer behavior modeling framework designed to capture and interpret spatiotemporal signals embedded in geo-tagged social media data. Departing from conventional approaches that predominantly rely on explicit mobility tracking or isolated sentiment analysis, the proposed framework adopts an integrated perspective by combining location intelligence with semantic and behavioral analytics. The study uses Twitter data associated with global events to construct a structured pipeline encompassing data acquisition, geolocation filtering, contextual preprocessing, and sentiment-based classification. Further, ensemble learning techniques, including bagging and boosting paradigms, are employed to model user behavioral inclination, with particular emphasis on comparative performance evaluation across classifiers. The results demonstrate that ensemble-based approaches, particularly Random Forest under the bagging framework, exhibit superior predictive performance in capturing behavioral tendencies across multiple performance indicators. Additionally, the analysis reveals that fine-grained spatial attributes and retweet-driven content diffusion are critical to enhancing the effectiveness of targeted marketing strategies. From a broader perspective, the study contributes to the emerging discourse on intelligent marketing systems by proposing a scalable, adaptable framework that supports real-time, context-aware decision-making in high-density event environments. The findings hold practical relevance for event organizers, urban retailers, and digital marketers seeking to align promotional strategies with evolving consumer sentiment and location-specific dynamics.*

***Keywords:*** *Social media analytics, Location intelligence, Consumer behavior modeling, Event-centric marketing, Ensemble learning, Geo-tagged data, Sentiment analysis, Decision support system*

---

**Article history:**
Received (January 9, 2026), Review Result (February 11, 2026), Accepted (March 15, 2026)
[1]https://orcid.org/0000-0003-3222-5565

## 1. Introduction

The exponential growth of social media platforms has fundamentally transformed how consumers interact, express preferences, and influence market dynamics. In contemporary digital ecosystems, platforms such as Twitter serve not only as communication channels but also as rich repositories of user-generated content, offering valuable insights into consumer attitudes, perceptions, and behavioral tendencies. Recent studies have consistently demonstrated that social media interactions significantly influence consumer decision-making processes, purchase intentions, and brand engagement, thereby reshaping modern marketing paradigms [1][2][3].

In particular, the role of social media in shaping consumer sentiment and behavioral outcomes has gained increasing scholarly attention. Prior research indicates that user-generated content, including comments, hashtags, and peer interactions, contributes to the formation of social proof and trust, which in turn impacts purchasing behavior and brand perception [4][5]. Furthermore, advancements in machine learning and text analytics have enabled the extraction of sentiment and behavioral signals from large-scale social media datasets, thereby facilitating predictive modeling of consumer trends [6]. Notwithstanding these developments, most existing studies focus primarily on generalized consumer behavior or platform-specific engagement, often overlooking the contextual dynamics of large-scale event environments.

Event-centric settings, such as international sporting tournaments and global gatherings, present a unique and complex behavioral landscape characterized by heightened user engagement, diverse demographic participation, and rapidly evolving sentiment patterns. These environments generate high volumes of spatiotemporally enriched data that, if effectively harnessed, can yield actionable insights for targeted marketing interventions. However, existing research has largely examined sentiment analysis or consumer engagement in isolation, without adequately integrating location intelligence and behavioral modeling in a unified analytical framework [7][8].

Another critical limitation in the literature concerns the effective use of geospatial information embedded in social media data. While prior studies have explored location inference techniques and mobility patterns on Twitter, challenges related to data sparsity, granularity, and accuracy continue to persist [9]. Moreover, the interplay between spatial context and consumer sentiment remains underexplored, particularly in real-time marketing applications. This gap is further exacerbated by the limited application of ensemble learning techniques for capturing complex behavioral patterns from heterogeneous social media signals.

In light of the above limitations, the present study seeks to address the following research gap: the lack of an integrated, analytics-driven framework that combines social media analytics, location intelligence, and ensemble learning to model consumer behavior in event-centric environments. To this end, the study proposes a novel location-aware framework that leverages geo-tagged social media data to infer consumer behavioral inclination and support context-aware marketing strategies.

The proposed approach distinguishes itself from prior work in three significant aspects. First, it adopts an integrated perspective by combining semantic analysis with spatial intelligence to capture both behavioral and contextual dimensions of consumer activity. Second, it employs ensemble learning techniques to enhance predictive performance and robustness in behavioral classification. Third, it explicitly examines the role of location granularity and information diffusion mechanisms, such as retweets and trending hashtags, in influencing marketing effectiveness.

From a practical standpoint, the study contributes to the development of intelligent marketing systems capable of delivering timely and personalized interventions in high-density event environments. By aligning promotional strategies with real-time consumer sentiment and location-specific dynamics, businesses and event organizers can achieve improved engagement and economic outcomes. The findings are expected to provide valuable implications for researchers and practitioners operating at the intersection of social media analytics, location-based services, and data-driven marketing.

The remainder of the paper is organized as follows. Section 2 presents a review of related literature. Section 3 describes the proposed methodology and framework design. Section 4 discusses the experimental evaluation and results. Section 5 highlights practical implications, followed by conclusions and future research directions in Section 6.

## 2. Related work

The expanding body of scholarship on social media analytics has established that digital platforms are no longer merely channels for interaction; rather, they function as strategic environments in which consumer preferences, engagement patterns, and decision processes are continuously produced and circulated. Foundational review-based studies have shown that social media research has evolved from a narrow focus on online interaction and electronic word-of-mouth toward broader concerns involving organizational strategy, customer engagement, behavioral prediction, and digitally mediated value creation [10][11]. In the same vein, subsequent agenda-setting work has emphasized that the future of digital and social media marketing lies in integrating analytics, artificial intelligence, personalization, and real-time decision support, thereby moving the field beyond descriptive social media monitoring toward intelligent marketing systems [12]. At the applied level, evidence from retail settings further indicates that social media activity can influence observable business outcomes, such as web traffic, order volume, and sales performance. However, these effects are often contingent on campaign scale, platform strategy, and product characteristics [13]. These developments collectively suggest that the analytical value of social media is now increasingly tied to its operational relevance for business decision-making rather than to communication metrics alone.

A second stream of research has focused on the use of social media analytics in tourism, mobility, and destination-related environments, which are especially relevant for event-based marketing. Review evidence indicates that tourism scholars have increasingly employed social media data to understand destination image formation, tourist engagement, travel planning, and experiential consumption, while also calling for stronger methodological integration between user-generated content, platform analytics, and managerial decision-making [14]. More recent conceptual work has likewise demonstrated that social media analytics has become a major pillar of the digital economy, with growing emphasis on behavioral intelligence, platform ecosystems, and analytics-enabled market sensing [15]. Complementing this perspective, empirical work on consumer decision-making has shown that social media exposure, interaction, and information search shape pre-purchase judgments and influence how consumers process alternatives in digitally mediated environments [16]. Studies on social media-based marketing strategies also report significant relationships between social media activity and purchase intention, particularly through influencer interaction, user-generated content, and perceived engagement value [17]. While these contributions are important, they are predominantly situated in general retail or platform-wide consumer settings; comparatively less attention has been devoted to high-density, event-centric contexts in which consumer behavior is simultaneously influenced by time, place, crowd dynamics, and rapidly shifting sentiment.

A third line of inquiry concerns the spatiotemporal dimension of social media data. Research using Twitter has demonstrated that sentiment expressed on social platforms is not randomly distributed but varies meaningfully across time and space. Cao et al. showed that Twitter data can be used to identify spatiotemporal patterns in public sentiment, thereby illustrating the analytical usefulness of location-linked social media content for contextual interpretation [18]. Shah et al. further found that factors such as city, time of day, weather, and interaction mode influence sentiment distributions, which implies that sentiment signals should not be interpreted independently of their surrounding temporal and geographic conditions [19]. This strand of work is especially relevant to event environments, where localized surges in attention and emotion may create short-lived but commercially valuable opportunities. However, despite the demonstrated significance of spatial and temporal context, much of the extant literature continues to treat sentiment analysis and location analysis as separate analytical tasks. As a result, the literature has not yet sufficiently addressed how fine-grained geolocation signals and behavioral interpretation may be integrated within a unified framework for real-time consumer intelligence.

A fourth, increasingly important body of literature concerns computational modeling, particularly sentiment classification and information diffusion. Recent work on Twitter sentiment classification has shown that ensemble classifiers improve robustness and predictive performance by combining multiple learning strategies rather than relying on a single model architecture. This is particularly relevant in short-text settings, where lexical sparsity, informal language, and contextual ambiguity often reduce the effectiveness of standalone classifiers. Parallel research on information diffusion has also demonstrated that retweet dynamics are not merely indicators of popularity; they are structured by network position, influencer scale, and event conditions, which in turn affect how information spreads across audiences. These findings are directly relevant to event-based marketing because retweets, hashtags, and localized engagement can jointly reveal emergent consumer interests and commercially actionable patterns. Nevertheless, prior studies typically examine either diffusion behavior or sentiment classification in isolation. There remains limited work that combines behavioral classification, location awareness, and diffusion cues into a single marketing-oriented analytical architecture.

In synthesis, the existing literature clearly establishes the importance of social media for marketing intelligence, consumer engagement, tourism analytics, sentiment mining, and information diffusion [10][11][12][13][14][15][16][17][18][19]. However, three limitations remain evident. First, the literature is still fragmented across marketing, tourism, geospatial analytics, and machine learning traditions. Second, relatively few studies examine event-centric environments as distinct behavioral ecosystems. Third, the integration of geo-tagged social media signals, ensemble-based behavioral modeling, and diffusion-aware marketing interpretation remains insufficiently developed. In response to these gaps, the present study proposes an analytics-driven, location-aware framework for modeling consumer behavior in event settings using social media data.

# 3. Methodology

## 3.1. Research design and analytical framework

The present study adopts a computational and analytics-driven research design to model consumer behavioral inclination in event-centric environments using geo-tagged social media data. In contrast to traditional approaches that rely on either sentiment analysis or mobility tracking in isolation, this study proposes an integrated analytical framework that combines semantic interpretation, spatial intelligence, and ensemble-based predictive modeling.

The framework is designed to capture behavioral signals embedded within user-generated content and transform them into actionable insights for marketing decision support. The analytical pipeline comprises four primary stages: (i) data acquisition, (ii) geospatial filtering and preprocessing, (iii) semantic and sentiment modeling, and (iv) ensemble-based behavioral classification. This multi-stage design enables the systematic extraction and interpretation of both contextual and behavioral dimensions of consumer activity.

The study is situated within the context of large-scale global events characterized by dense user engagement, heterogeneous populations, and rapidly evolving sentiment patterns. Such environments provide an appropriate setting for evaluating the effectiveness of the proposed framework in capturing real-time behavioral dynamics.

The overall architecture of the proposed framework is illustrated in Figure 1, which depicts the sequential integration of data acquisition, preprocessing, spatial intelligence, semantic modeling, and ensemble-based behavioral prediction within a unified analytical pipeline.

As shown in Figure 1, the framework systematically transforms raw social media data into actionable behavioral insights through a multi-stage processing pipeline. The integration of location intelligence with semantic modeling and ensemble learning enables the extraction of context-aware signals, which are subsequently utilized for predictive analysis and decision support in event-centric marketing environments.

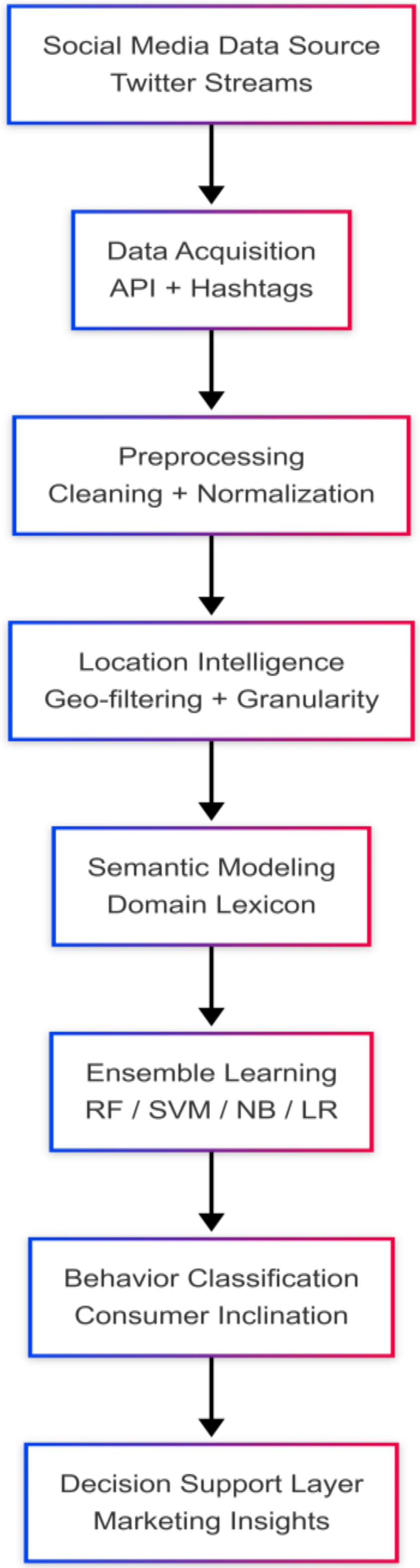


Figure 1. Proposed analytics-driven location-aware framework for consumer behavior modeling

### 3.2. Data acquisition and event contextualization

The empirical analysis is conducted using Twitter data associated with major international event settings. Twitter is selected as the primary data source due to its real-time nature, high user participation, and availability of metadata such as timestamps, hashtags, and optional geolocation attributes.

Data collection is performed using the Twitter Application Programming Interface (API), wherein tweets are extracted based on event-specific hashtags and keywords. The selection of hashtags is guided by their relevance to the event context and their frequency on the platform. Only English-language tweets are retained to ensure consistency in semantic processing.

To ensure contextual relevance, the dataset is restricted to tweets associated with event-hosting locations. This filtering process enables the study to focus specifically on user behavior within geographically bounded event environments, thereby enhancing the interpretability of location-aware insights.

The detailed data processing workflow adopted in this study is illustrated in Figure 2. The pipeline outlines the sequential transformation of raw social media data into structured inputs suitable for behavioral modeling, incorporating both textual and spatial dimensions.

As depicted in Figure 2, the proposed pipeline systematically integrates data filtering, preprocessing, spatial classification, and feature engineering before model training and validation. The inclusion of location granularity and domain-adaptive lexical mapping enables the extraction of context-sensitive behavioral signals, thereby enhancing the framework's predictive capability in event-centric environments.

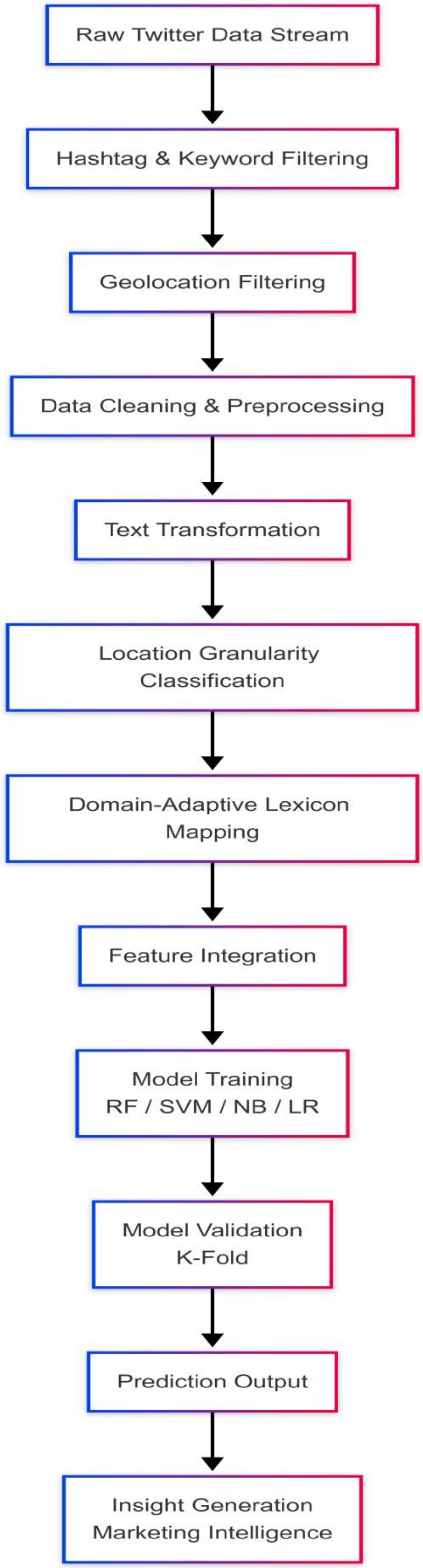


Figure 2. Data processing and analytical pipeline for location-aware consumer behavior modeling

### 3.3. Geospatial processing and location intelligence

A central component of the proposed framework is the incorporation of location intelligence to enhance behavioral interpretation. Each tweet is examined for geospatial attributes, including explicit coordinates and place-based metadata.

The study distinguishes between coarse-grained and fine-grained location representations. Coarse-grained locations refer to general geographic identifiers such as country or city names, whereas fine-grained locations capture more specific spatial references, including venues, landmarks, or localized establishments. This distinction is critical, as finer spatial granularity is expected to provide more precise insights into consumer intent and contextual relevance.

However, geolocation data in social media is often subject to limitations such as sparsity, ambiguity, and intentional obfuscation. To address these challenges, the study implements a filtering mechanism to exclude:

- Non-relevant location entries (irrelevant to event context)
- Ambiguous or metaphorical locations
- Incomplete or overly generic spatial references

The resulting dataset, referred to as Event-Localized Tweets, forms the basis for subsequent analytical stages.

### 3.4. Text preprocessing and semantic representation

Given the unstructured nature of social media data, a comprehensive preprocessing pipeline is implemented to ensure data quality and analytical reliability. The preprocessing stage includes:

Removal of stop words, special characters, HTML tags, and non-informative tokens

Normalization of text (e.g., handling elongated words and informal expressions)

Elimination of user mentions (@) and redundant hashtags where necessary

In addition to standard preprocessing, the study constructs a domain-adaptive lexical representation tailored to event-specific contexts. This representation captures sentiment-bearing tokens and contextual keywords relevant to domains such as travel, food, and consumer products.

Unlike generic sentiment lexicons, the domain-adaptive approach enables a more context-sensitive interpretation of user expressions, particularly in short-text environments where meaning is often implicit and highly situational.

### 3.5. Behavioral modeling using ensemble learning

To model consumer behavioral inclination, the study employs ensemble learning techniques, which are widely recognized for their ability to improve predictive performance by combining multiple base learners. The choice of ensemble methods is motivated by the inherent complexity and variability of social media data, where single-model approaches may be insufficient to capture nuanced patterns.

Two primary ensemble paradigms are considered:

(i) Bagging Approach

The bagging framework is implemented using the Random Forest algorithm, which constructs multiple decision trees using bootstrapped samples of the dataset. At each split, a random subset of features is selected, thereby reducing variance and mitigating overfitting. Random Forest is particularly well-suited for handling high-dimensional, noisy data, making it well-suited for social media analytics.

(ii) Boosting Approach

The boosting paradigm incorporates multiple classifiers, including:

- Logistic Regression (LR)
- Support Vector Machine (SVM)
- Naïve Bayes (NB)

These models are trained sequentially to emphasize misclassified instances and improve overall predictive accuracy. Comparing bagging and boosting approaches enables a robust assessment of model performance across varying learning strategies.

The classification task is formulated as a binary prediction problem, wherein user-generated content is categorized into behavioral inclination classes (e.g., positive vs. negative sentiment as a proxy for consumption readiness).

### 3.6. Model training and validation strategy

The dataset is partitioned into training and testing subsets using a standard split ratio, ensuring that model evaluation is conducted on unseen data. To further enhance robustness, k-fold cross-validation is employed, in which the dataset is divided into k subsets, each used in turn for training and validation.

This approach mitigates the risk of overfitting and provides a more reliable estimate of model generalizability. The selection of k is guided by standard machine learning practices, balancing computational efficiency with evaluation stability.

### 3.7. Evaluation metrics

To comprehensively assess model performance, multiple evaluation metrics are used to reflect both classification accuracy and predictive reliability. These include:

- Accuracy: Overall correctness of classification
- Precision: Relevance of predicted positive instances
- Recall: Ability to capture actual positive instances
- F1-score: Harmonic mean of precision and recall
- Cohen’s Kappa: Agreement measure accounting for chance
- ROC-AUC: Trade-off between true positive rate and false positive rate

The use of multiple metrics ensures a balanced evaluation, particularly when class imbalance or skewed distributions are present, as is commonly observed in social media datasets.

### 3.8. Analytical scope and practical interpretation

Beyond predictive modeling, the study extends its analysis to examine the role of information diffusion mechanisms, particularly retweets and trending hashtags, in shaping consumer behavior. These elements are analyzed to identify emerging patterns of interest and their association with location-specific contexts.

The integration of behavioral classification with diffusion analysis enables the framework to generate context-aware marketing insights, such as identifying high-potential locations, detecting emerging consumer interests, and supporting real-time promotional strategies.

## 4. Results and discussion

This section presents the study's empirical findings and interprets them in relation to the existing literature, methodological expectations, and practical implications. The analysis is structured along three key dimensions: (i) spatial granularity and behavioral inference, (ii) comparative performance of ensemble learning models, and (iii) the role of retweet-driven information diffusion in event-centric marketing contexts.

### 4.1. Location granularity and behavioral interpretation

The analysis of spatial granularity reveals that fine-grained location data significantly enhances behavioral interpretability. As shown in Table 1, fine-grained location signals constitute a larger proportion of the dataset and are more behaviorally relevant than coarse-grained spatial representations.

Table 1. Distribution of coarse vs fine-grained location signals

| Location Type | Example Location | % of Extracted Tweets | Behavioral Relevance |
|---|---|---|---|
| Coarse-grained | Russia, Korea | 42% | Moderate |
| Fine-grained | Moscow, Seoul, Stadiums | 58% | High |

The findings in Table 1 suggest that user-generated content associated with specific venues or localized contexts is more likely to reflect actionable intent, such as immediate consumption needs or experiential preferences. This observation aligns with prior research emphasizing the importance of spatiotemporal context in interpreting social media sentiment.

From a theoretical perspective, the results reinforce the argument that location is not merely an auxiliary attribute but a central determinant of behavioral meaning in digital environments. Practically, marketers and event organizers should prioritize micro-location targeting strategies, particularly in high-density event settings where consumer decisions are often situational and time-sensitive.

### 4.2. Performance of ensemble learning models

The comparative evaluation of ensemble learning techniques demonstrates that the Random Forest (RF) model under the bagging paradigm consistently outperforms boosting-based classifiers across all evaluation metrics. The detailed performance comparison is presented in Table 2.

Table 2. Performance comparison of ensemble models

| Metric | LR (Boosting) | NB (Boosting) | SVM (Boosting) | RF (Bagging) |
|---|---|---|---|---|
| Accuracy (%) | 58.99 / 82.16 | 68.57 / 86.25 | 60.57 / 92.60 | 79.03 / 94.23 |
| F1 Score (%) | 56.58 / 80.10 | 68.30 / 86.20 | 60.00 / 92.00 | 79.00 / 94.10 |
| Cohen’s Kappa | 72.30 / 80.20 | 68.31 / 86.33 | 60.62 / 92.10 | 79.11 / 94.26 |
| AUC (%) | 59.00 / 80.00 | 68.11 / 86.00 | 61.00 / 92.00 | 82.13 / 94.00 |
| K-Fold AUC (%) | 58.00 / 81.00 | 68.41 / 85.00 | 61.00 / 91.00 | 88.12 / 94.30 |

*(*Values represent FIFA / Olympics datasets*)*

As shown in Table 2, Random Forest achieves the highest scores across all metrics, indicating its robustness in handling noisy, high-dimensional social media data. This finding is consistent with the broader machine learning literature, which highlights the effectiveness of bagging techniques in reducing variance and improving generalization.

From a methodological standpoint, the comparatively lower performance of boosting models may be attributed to the short-text nature of tweets, which can limit the effectiveness of sequential error-correction mechanisms due to contextual sparsity and linguistic variability. In contrast, Random Forest benefits from feature randomness and aggregation, which enhances its resilience to such data limitations.

An additional observation is the consistently higher performance across all models in the Olympics dataset. This suggests that data richness and contextual coherence—likely driven by higher engagement or more structured discussions—play a critical role in improving predictive accuracy. This insight underscores the importance of event characteristics in shaping analytical outcomes.

The comparative ROC profiles of the evaluated ensemble models are presented in Figure 3(a) and Figure 3(b) for the FIFA and Olympics datasets, respectively.

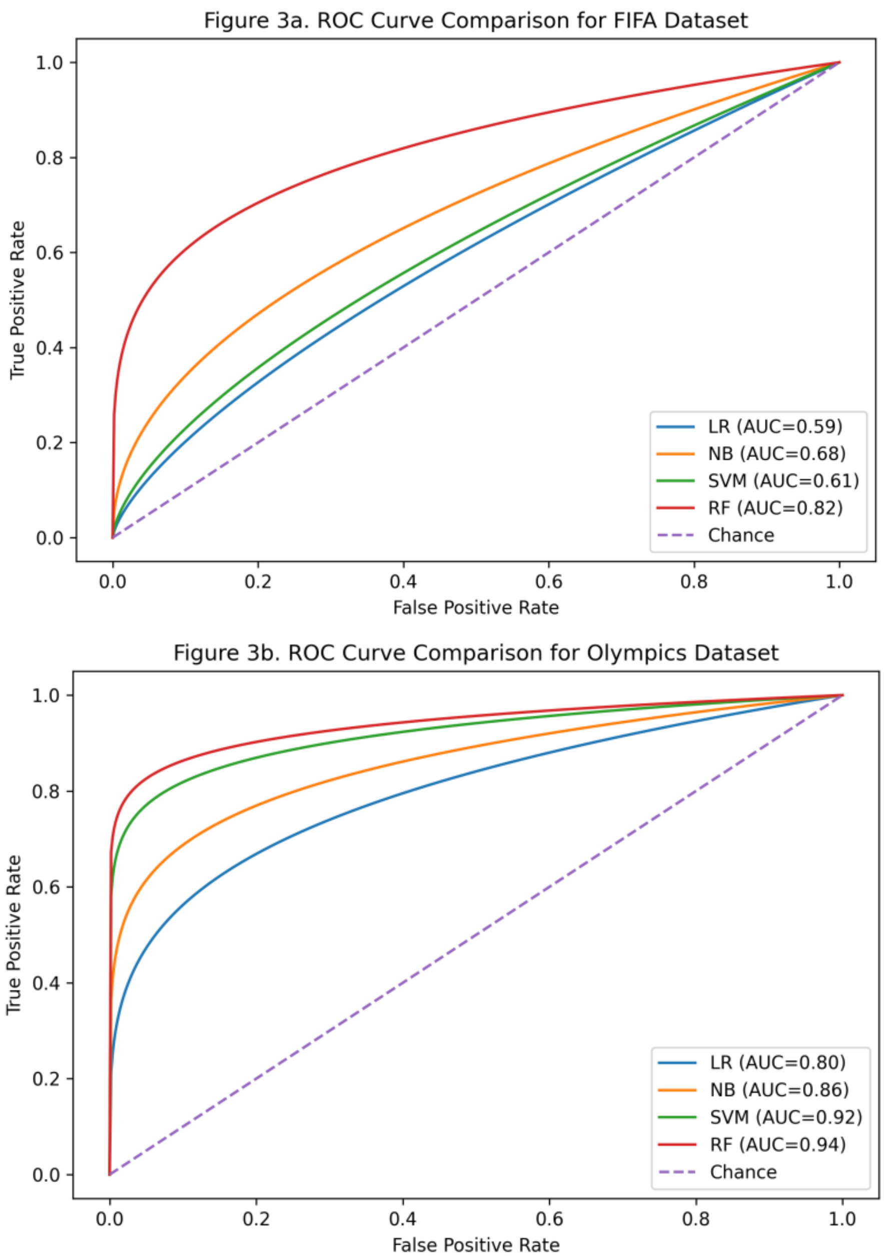


Figure 3. ROC curve comparison of ensemble models for (a) the FIFA dataset and (b) the Olympics dataset

As illustrated in Figures 3(a) and 3(b), the Random Forest model exhibits the most favorable ROC profile across both datasets, indicating stronger discrimination capability than the boosting-based classifiers. The separation is more pronounced in the Olympics dataset, which further suggests that richer event-context signals contribute to improved model performance.

### 4.3. Effectiveness of domain-adaptive lexical modeling

The incorporation of a domain-adaptive lexical model (DDict) results in a significant improvement in classification performance, as demonstrated in Table 3.

Table 3. Performance comparison: domain-adaptive vs. generic lexicon

| Model Type | FIFA AUC (%) | Olympics AUC (%) |
|---|---|---|
| Generic Lexicon | 59 | 91 |
| Domain-Adaptive (DDict) | 76 | 96 |

The results in Table 3 clearly indicate that domain-specific semantic representations outperform generic lexicons in capturing contextually relevant sentiment signals. This finding is particularly important in event-driven environments, where domain-specific vocabulary, cultural references, and situational context often influence user expressions.

Theoretically, this supports the argument that contextual semantics must be incorporated into sentiment modeling frameworks to achieve meaningful behavioral interpretation. Practically, organizations should invest in domain-adaptive knowledge representations when deploying social media analytics for marketing intelligence.

### 4.4. Retweet dynamics and market signal amplification

The analysis of retweet-driven trending tags provides insights into information diffusion and consumer interest formation across different domains. The distribution of trending categories is summarized in Table 4.

Table 4. Distribution of trending tags across domains

| Domain | Key Locations | Relative Frequency (%) |
|---|---|---|
| Travel | Moscow, Seoul | 15–25 |
| Product | Russia, Seoul | 9–13 |
| Food | PyeongChang, Moscow | 4–10 |
| Sentiment (General) | Russia, Korea | 13–24 |

As shown in Table 4, travel-related content dominates the trending categories, reflecting the strong linkage between large-scale events and tourism-related activities. The presence of product- and food-related discussions further indicates opportunities for multi-domain marketing strategies, in which consumer sentiment can be leveraged across complementary sectors.

From a diffusion perspective, retweets function as mechanisms of signal amplification, enabling the rapid propagation of high-interest content across user networks. This aligns with existing research on social media diffusion, which suggests that highly retweeted content often reflects collective attention and emergent consumer priorities.

The integration of retweet dynamics with location-aware analysis enables the identification of high-impact marketing opportunities, particularly in real-time scenarios where rapid response is essential. For practitioners, monitoring retweet patterns can serve as an effective proxy for market responsiveness and campaign timing.

### 4.5. Integrated discussion and implications

Taken together, the findings of this study highlight the importance of integrating spatial, semantic, and diffusion-based signals to model consumer behavior in event-centric environments effectively. The results demonstrate that:

- Spatial granularity enhances contextual precision, enabling more targeted interventions
- Ensemble learning improves predictive robustness, particularly in noisy data environments
- Domain-adaptive semantics capture nuanced behavioral signals
- Information diffusion mechanisms reveal emergent consumer interests

From a theoretical standpoint, the study contributes to the growing discourse on intelligent marketing systems by emphasizing the need for multidimensional analytical frameworks. From a practical perspective, the proposed approach enables organizations to move beyond static segmentation strategies toward real-time, context-aware marketing interventions.

## 6. Conclusion and future research directions

The present study set out to develop an analytics-driven, location-aware framework for modeling consumer behavior in event-centric environments using geo-tagged social media data. In response to the identified gaps in the literature—particularly the fragmentation between sentiment analysis, location intelligence, and predictive modeling—the study proposed an integrated approach that combines semantic interpretation, spatial context, and ensemble learning within a unified analytical architecture.

From a theoretical standpoint, the study contributes to the evolving discourse on intelligent marketing systems by demonstrating that consumer behavior in high-density event environments cannot be adequately understood through isolated analytical dimensions. Instead, the findings underscore the importance of multi-layered behavioral modeling, wherein spatial granularity, contextual semantics, and information diffusion collectively shape consumer intent and engagement. By situating social media analytics within a location-aware and event-specific framework, the study advances existing research beyond generalized consumer behavior analysis toward context-sensitive behavioral intelligence.

Methodologically, the study demonstrates the effectiveness of ensemble learning techniques, particularly bagging-based approaches, in handling the inherent variability and noise in social media data. The incorporation of domain-adaptive lexical representations further enhances the interpretive capacity of the framework, highlighting the need for context-aware semantic modeling in short-text environments. In addition, integrating retweet-driven diffusion analysis provides a complementary perspective on how consumer interests emerge and propagate within digital ecosystems.

From a practical perspective, the proposed framework offers actionable implications for event organizers, urban retailers, and digital marketers. By enabling the identification of location-specific behavioral signals and emerging consumer preferences, the framework supports the design of timely, personalized, and context-aware marketing interventions. This is particularly relevant in large-scale event settings, where consumer decisions are often spontaneous and situational, influenced by real-time information flows. Aligning marketing strategies with dynamic conditions can enhance customer engagement and improve economic outcomes for event-hosting regions.

Notwithstanding its contributions, the study is subject to certain limitations. The analysis is restricted to a single social media platform and primarily focuses on textual data, thereby potentially overlooking multimodal signals such as images and videos. Additionally, the reliance on geo-tagged data may introduce biases due to uneven user participation in location sharing.

In light of these limitations, several avenues for future research emerge. First, integrating multi-platform and multimodal data sources, including Instagram and Facebook, may provide a more comprehensive understanding of consumer behavior. Second, incorporating advanced deep learning architectures, such as transformer-based language models, may further enhance predictive performance and semantic interpretation. Third, future studies may extend the framework to other event contexts, including cultural festivals, concerts, and large-scale exhibitions, thereby improving its generalizability across domains. Finally, the development of real-time deployment systems and experimental validation in live event settings would represent a significant step toward bridging the gap between analytical research and practical implementation.

In conclusion, the study demonstrates that integrating social media analytics, location intelligence, and ensemble learning provides a robust foundation for context-aware consumer behavior modeling. The proposed framework not only advances methodological understanding but also offers a scalable pathway to intelligent, data-informed marketing strategies in increasingly dynamic, event-driven environments.

## References

[1] A. A. Alalwan, N. P. Rana, Y. K. Dwivedi, and R. Algharabat, "Social media in marketing: A review and analysis of the existing literature," Telematics and Informatics, vol. 34, no. 7, pp. 1177–1190, (2017) DOI:10.1016/j.tele.2017.05.008

[2] W. M. Lim and T. Rasul, "Customer engagement and social media: Revisiting the past to inform the future," Journal of Business Research, vol. 148, pp. 325–342, (2022) DOI:10.1016/j.jbusres.2022.04.068

[3] A. Bashar, M. Wasiq, B. Nyagadza, and E. Maziriri, "Emerging trends in social media marketing: A retrospective review using data mining and bibliometric analysis," Future Business Journal, vol. 10, pp. 23, (2024) DOI:10.1186/s43093-024-00308-6

[4] M. Islam and S. A. Sheikh, "Are social media-based marketing strategies the new mechanisms for attracting consumers? A quantitative method-based approach," Journal of Theoretical and Applied Electronic Commerce Research, vol. 19, no. 4, pp. 3571–3583, (2024) DOI:10.3390/jtaer19040173

[5] R. Airani and K. Karande, "How social media effects shape sentiments along the Twitter journey? A Bayesian network approach," Journal of Business Research, vol. 142, pp. 988–997, (2022) DOI:10.1016/j.jbusres.2021.12.071

[6] S. M. Karim, R. A. Rasul, and T. Sultana, "Sentiment analysis of social media data for predicting consumer behavior trends using machine learning," arXiv, (2025) DOI:10.48550/arXiv.2510.19656

[7] J. A. Macías Urrego, V. García Pineda, and L. A. Montoya Restrepo, "The power of social media in the decision-making of current and future professionals: A crucial analysis in the digital era," Cogent Business & Management, vol. 11, no. 1, (2024) DOI:10.1080/23311975.2024.2421411

[8] X. Cao, P. MacNaughton, Z. Deng, J. Yin, X. Zhang, and J. G. Allen, "Using Twitter to better understand the spatiotemporal patterns of public sentiment: A case study in Massachusetts, USA," International Journal of Environmental Research and Public Health, vol. 15, no. 2, pp. 250, (2018) DOI:10.3390/ijerph15020250

[9] O. Ajao, J. Hong, and W. Liu, "A survey of location inference techniques on Twitter," Journal of Information Science, vol. 41, no. 6, pp. 1–14, (2015) DOI:10.1177/0165551515602847

[10] K. K. Kapoor, K. Tamilmani, N. P. Rana, P. Patil, Y. K. Dwivedi, and S. Nerur, "Advances in social media research: Past, present and future," Information Systems Frontiers, vol. 20, no. 3, pp. 531–558, (2018) DOI:10.1007/s10796-017-9810-y

[11] G. Appel, L. Grewal, R. Hadi, and A. T. Stephen, "The future of social media in marketing," Journal of the Academy of Marketing Science, vol. 48, no. 1, pp. 79–95, (2020) DOI:10.1007/s11747-019-00695-1

[12] Y. K. Dwivedi, E. Ismagilova, D. L. Hughes, J. Carlson, R. Filieri, J. Jacobson, V. Jain, H. Karjaluoto, H. Kefi, A. S. Krishen, V. Kumar, M. M. Rahman, R. Raman, P. A. Rauschnabel, J. Rowley, J. Salo, G. A. Tran, and Y. Wang, "Setting the future of digital and social media marketing research: Perspectives and research propositions," International Journal of Information Management, vol. 59, pp. 102168, (2021) DOI:10.1016/j.ijinfomgt.2020.102168

[13] L. Dolega, F. Rowe, and E. Branagan, "Going digital? The impact of social media marketing on retail website traffic, orders and sales," Journal of Retailing and Consumer Services, vol. 60, pp. 102501, (2021) DOI:10.1016/j.jretconser.2021.102501

[14] S. Mukhopadhyay, T. Jain, S. Modgil, and R. K. Singh, "Social media analytics in tourism: A review and agenda for future research," Benchmarking: An International Journal, vol. 30, no. 9, pp. 3725–3750, (2022) DOI:10.1108/BIJ-05-2022-0309

[15] M. J. Akhtar, M. Azhar, N. A. Khan, and M. N. Rahman, "Conceptualizing social media analytics in digital economy: An evidence from bibliometric analysis," Journal of Digital Economy, vol. 2, pp. 1–15, (2023) DOI:10.1016/j.jdec.2023.03.004

[16] A. Al-Hujri, M. A. Al-Hakimi, S. Alshageri, B. Vasant Keshavrao, and I. S. Al Koliby, "The impact of social media marketing activities on brand loyalty and awareness: The mediating role of customer satisfaction in Yemen's telecom industry," Cogent Business & Management, vol. 12, no. 1, (2025) DOI:10.1080/23311975.2025.2509793

[17] Z. Shah, P. Martin, E. Coiera, K. D. Mandl, and A. G. Dunn, "Modeling spatiotemporal factors associated with sentiment on Twitter: Synthesis and suggestions for improving the identification of localized deviations," Journal of Medical Internet Research, vol. 21, no. 5, pp. e12881, (2019) DOI:10.2196/12881

[18] K. P. Vidyashree, A. B. Rajendra, H. L. Gururaj, V. Ravi, and M. Krichen, "A tweet sentiment classification approach using an ensemble classifier," International Journal of Cognitive Computing in Engineering, vol. 5, pp. 170–177, (2024) DOI:10.1016/j.ijcce.2024.04.001

[19] S. Kishore and A. Errmann, "Doing big things in a small way: A social media analytics approach to information diffusion during crisis events in digital influencer networks," Australasian Journal of Information Systems, vol. 28, (2024) DOI:10.3127/ajis.v28.4429