Predicting Visitors' Digital Footprints Using Spatial Features and Synthetic Pedestrian Data

Müslüm Hacar(✉)

Yildiz Technical University, Esenler, Istanbul 34210, Türkiye
mhacar@yildiz.edu.tr

Abstract. Assessing digital footprints of visitors has often depended on a pedestrian count data representing the human density in urban environment. The data is commonly generated via sensor, cellular phone network, or CCTV technologies, which involves high costs and raises concerns regarding personal data protection. This study presents a novel approach for predicting visitor ratings of eating and drinking out venues in Bakırköy, Istanbul, without relying on a real pedestrian counting system. In our work, a synthetic pedestrian count measure is generated by interpolating the overall review volume data extracted from a complete set of points-of-interest data via Google Maps API. This measure is aggregated to streets, and additional spatial measures are used to further describe the urban environment. Some street geometric features (i.e., sinuosity and street length) that showed limited predictive value in earlier studies were removed from the analysis. A random forest machine learning model was applied to predict visitors' high/low ratings, achieving a 76% F1-score. The results suggest that digital traces left by visitors can be effectively predicted through alternative spatial measures and synthetic data. The proposed approach provides a cost-effective and privacy-respecting method that can support business decision-making in the context of digital consumer behavior.

Keywords: Synthetic Pedestrian Count · Digital Footprint · Urban Analytics · Street Network · Point of Interest

1 Introduction

Digital records left by visitors play a significant role in understanding urban life. In recent years, the analysis of digital footprints, which refers to the records of online behavior that individuals leave behind, including reviews and rating, has become a popular method for evaluating consumer behavior [1–6]. These records offer valuable information about consumer satisfaction and service quality, which can be used by urban planners and business stakeholders [3–6]. In particular, eating and drinking out (EDO) venues generate a considerable amount of digital data that can be used to analyze consumer satisfaction and guide planning decisions. In recent years, studies described below have shown that digital records provide an effective measure for assessing urban dynamics and visitor behavior. This section follows with (1) a literature review part on digital footprint research

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2026
O. Gervasi et al. (Eds.): ICCSA 2025 Workshops, LNCS 15897, pp. 391–402, 2026.
https://doi.org/10.1007/978-3-031-97660-5_26

studies, specifically in urban areas, then (2) describes the gaps in the current analysis, and (3) concludes with the motivations to propose a novel approach to generating synthetic pedestrian data which can be used as a feature for the prediction of visitor ratings.

The use of digital footprints for urban analysis is supported by several studies. Valanarasu [1] examined the potential of the footprint for personality prediction. In his study, he investigated the growing use of social media digital footprints to predict user personality traits by analyzing publicly accessible posts and status updates from multiple platforms, including Facebook, Twitter, and YouTube. His work introduced a novel hybrid machine learning approach that integrates dynamic multi-context information to achieve higher accuracy in personality prediction. Adikari et al. [2] proposed an approach that leverages Natural Language Processing to transform unstructured hotel reviews into a structured representation of emotions and hotel attributes. Their method, validated on 22,450 reviews from 44 hotels, enables segmentation of customer and hotel profiles to support targeted management strategies and informed decision-making. Di Clemente et al. [5] introduced a framework that applies text compression techniques on credit card transaction sequences, capitalizing on Zipf-like distributions to reveal significant purchasing patterns. Their method successfully clusters consumers into groups with similar demographic and behavioral characteristics based on the temporal order of their transactions. Khan and Loan [6] examined how public library users in Delhi utilize Google Maps to share opinions about facilities and services, using content and sentiment analysis to classify the reviews. Their findings indicate that the platform serves as an effective interactive feedback system, with the majority of reviews expressing positive sentiments while highlighting specific areas for improvement. Hacar et al. [7] investigated how urban spatial configurations, including street prominence and place type, influence the digital footprints captured through online reviews and ratings. By applying Space Syntax metrics to Google Maps points of interest (POIs) in Sassi di Matera, Italy, they demonstrated that the relationship between urban form and visitor review behavior varies significantly depending on the type of place. Muhammed et al. [8] developed a conceptual framework through a systematic literature review to identify the key factors that influence customers' willingness to leave digital footprints on social media. Their study finds that intrinsic psychological traits, technological convenience, social influences, and concerns about privacy and security are critical determinants of this behavior. Girardin et al. [9] presented innovative methods for analyzing spatiotemporal data derived from user-generated sources like cell phone networks and georeferenced photos. Their approach effectively reveals tourist presence and movement patterns, providing valuable indicators of urban dynamics. Martí et al. [10] proposed a reference framework that uses volunteered user-generated content from multiple location-based social networks to identify and characterize tourist activity centers in urban settings. Their method enables a fine-grained analysis of tourism dynamics, thereby offering a complementary tool for urban planning decision-making. Jia and Liu [11] investigated the roles of online review valence and usefulness in influencing consumer behavior, suggesting that the positivity of a review offers more diagnostic value than its usefulness rating. Their mixed-method research, which combines large-scale panel data with experimental studies, demonstrates that highly useful reviews increase consumer adoption only when perceived as positive. Li and Hecht [12] investigated discrepancies between restaurant

ratings on Google Maps and Yelp, revealing that ratings on Google Maps are on average 0.7 stars higher, particularly for chain restaurants. Their study also shows minimal overlap in top-ranked lists across platforms, questioning the reliability of using a single review source for local searches. Akkaya et al. [13] proposed a machine learning method that uses Google Maps Reviews to extract structured insights on park features and social structures in Istanbul. This approach offers a practical alternative to traditional, labor-intensive public feedback collection methods, supporting urban planning decisions. Owour and Hochmair [14] reviewed 110 social media apps to evaluate their potential for geospatial research, highlighting that only a fraction provided robust APIs and relevant spatial data. Their findings suggest that several lesser-known apps merit further exploration, as they could uncover underexplored aspects of urban dynamics. Tang et al. [15] applied deep learning techniques to segment panoramic urban images and analyzed how built environment attributes affect restaurant popularity. Their results indicate that pedestrian flow is the most critical factor influencing restaurant ratings, while higher densities of traffic signs negatively correlate with review counts. Rajabi et al. [16] assessed healthcare center accessibility in Tehran using Google POI data and clustering algorithms, revealing a significant imbalance in service distribution across the city. Their analysis underscores a misalignment between healthcare resources and population density, emphasizing the need for targeted urban planning. Tse et al. [17] examined Canadian acute care hospitals by correlating Google Reviews with hospital characteristics and patient experience scores, uncovering substantial variability among hospitals. Their study demonstrates that online reviews can serve as a readily accessible, real-time tool for monitoring and improving patient care. Leiras and Eusébio [18] analyzed over a thousand online traveler reviews from Aveiro and A Coruña to assess the perceived image of Accessible Tourism Destinations. Their findings revealed that key accessibility features, such as parking availability and disability adaptations, significantly impact visitor satisfaction, offering a strategic framework for tourism management. Öztürk Hacar et al. [19] analyzed how urban spatial characteristics, such as pedestrian counts, residential density, and street geometry, influence digital consumer behavior at EDO places. Their random forest model applied to Google Maps reviews in Melbourne's city center indicates that pedestrian counts and residential density are strong predictors of review volumes and visitor ratings, while measures of street centrality and geometry exhibit varied effects, suggesting a promising alternative for urban planning and business decision-making.

Many municipalities and private enterprises are interested in using digital footprints for strategic planning and decision-making. Assessing digital footprints of visitors has often depended on pedestrian count data which are commonly generated via sensor, cellular phone network, or CCTV technologies, involving high costs and raising concerns regarding personal data protection. Instead of using a count data, using a predicted pedestrian volume can play a critical role in reducing those disadvantages. Jiang et al. [20] developed a spatiotemporal machine learning model to estimate pedestrian volume in Melbourne. On the other hand, in contexts where historical sensor data are unavailable, digital footprints from location-based platforms can serve as a valuable proxy for modeling pedestrian activity. This paper proposes an approach to generating synthetic pedestrian data derived from Google POI review volumes and supportive spatial features

to predict the visitor ratings at EDO places in Bakirköy, Istanbul. By avoiding the use of direct and indirect pedestrian counting technology, the proposed approach addresses key issues related to cost and privacy while offering businesses a practical framework for digital analytics.

2 Materials and Methods

2.1 Study Area and Data Structures

The study is conducted in Bakırköy, a prominent center on Istanbul's European side, particularly renowned for its dense clustering of university preparatory, post-graduate courses, and shopping facilities (see Fig. 1). This concentration of mobility has resulted in a high volume of pedestrian movement, which in turn has spurred the emergence of numerous alternative eating and drinking venues to meet the dynamic needs of the population.

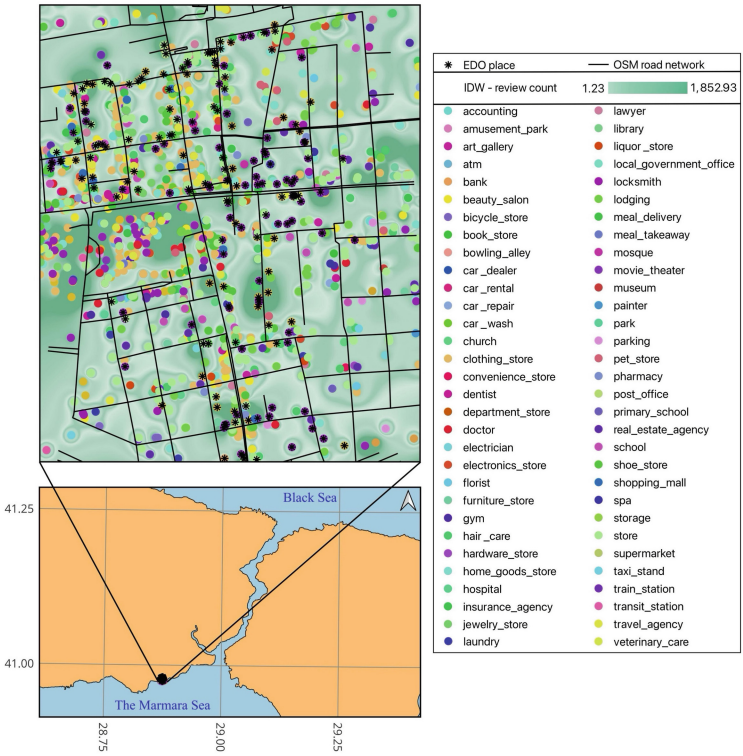

Fig. 1. Study area and the data representations.

Data for this study are primarily sourced from POI points obtained via the Google Maps API [21] and OSM road networks downloaded via OSM Export [22]. The POI

records include essential attributes such as review count, rating, and type, which are utilized as the key features in the analysis. In particular, the rating attribute is categorized into two classes: High and Low, using a quantile approach. These categories serve as the dependent variable for predicting visitors' digital footprints. Moreover, OSM road network offers a comprehensive representation of urban transportation. In this study, an arc-node topology of the road network was constructed by first applying a dissolve operation followed by split operations by each junction. All road-based features are derived from the road network topology.

2.2 The Proposed Approach to Feature Generation and Prediction Model

Google POIs data is specifically for EDO establishments, which provides semantic details like place type, review count, and rating [21]. To ensure the reliability of our analysis on visitor behavior, we included only the places that have accumulated at least 10 reviews. This threshold helps us concentrate on consumer-focused locations, facilitating a deeper investigation into the correlation between the spatial distribution of places and consumers' digital footprints.

Instead of employing direct pedestrian count data, typically gathered from sensors, cellular networks, or CCTV, review count data of all types of POIs is used as a surrogate metric to indicate the intensity of consumer presence across the urban landscape. Additionally, to prepare the count data for the following spatial analysis, we applied an Inverse Distance Weighting (IDW) interpolation in a QGIS (version 3.28) environment, which allows to generate continuous surfaces that approximate pedestrian density patterns [23]. Unlike Kriging, IDW does not require complex statistical assumptions or parameters such as spatial autocorrelation or variogram models. IDW assumes that the influence of each data point decreases with distance, which fits well with the idea that areas near popular places with high review counts are likely to have more pedestrian activity. Each grid point's weighted value on the surface is associated to its closest street. Thus, a mean value of several grid points, which is hereafter referred as synthetic-PC, is assigned to each street (see Fig. 2).

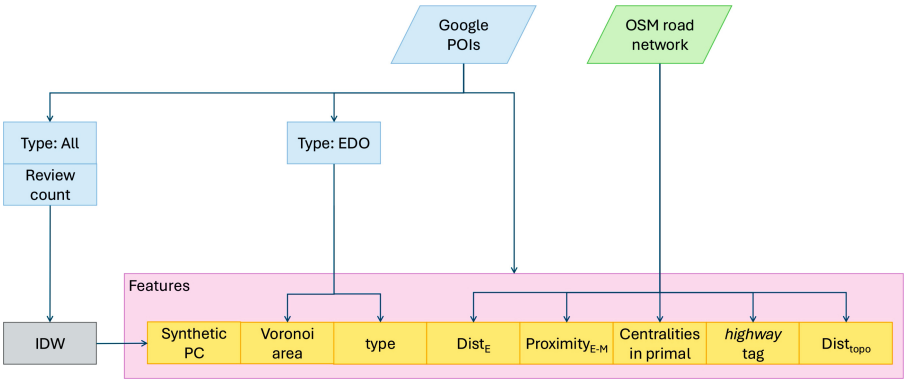

Fig. 2. Workflow of the feature generation.

The review counts at the places within the shopping malls (i.e., Carousel and Capacity shopping malls) are included in the synthetic-PC process to represent the crowd pedestrian population. However, the places are excluded in the target EDO establishment list since they are relatively independent from the neighboring streets but more depended on the indoor locations inside the building of shopping mall.

Voronoi polygons are created to express local proximity as a feature for each place. The area of each polygon is assigned to the regarding central node, POI. Also, the type of EDO POIs (i.e., bakery, bar, café, and restaurant) is used as a feature expressing the diversity of source data and is considered that it may represent the unique properties among others.

OSM road network data includes *highway* tags that classify roads based on their functional type [24]. These types are assumed to have different influences on visitor behavior and accessibility to services. To further quantify the connectivity and importance of individual street segments, centrality measures using a primal graph representation through the Python NetworkX package (version 3.2.1) are computed [25]. Eight metrics to reveal the representation of connectivity within the network are used: betweenness (B), closeness (C), degree (D), eigenvector (E), harmonic (H), Katz centrality (K), load (L), and PageRank (P). These metrics help explain how integrated or isolated each place is within the urban grid [19].

Rest of the features are generated by using both Google POIs and OSM. Firstly, $Dist_E$ is the Euclidean distance between the place and the closest node which has the degree of connectivity as more than two. It is considered to detect the closest potentially densified junctions. Secondly, $Proximity_{E\text{-}M}$ is the binary code referring whether the place is close to an end node (E) or middle node (M) of the closest street (see Fig. 3). This measure helps to depict a significant proximity information about an EDO place if it is more like around the junction of a street or not. Thirdly, $Dist_{topo}$ is the topological distance, which gives total number of steps over the street network from an EDO place to the closest pedestrian-friendly roads (i.e., *footway*, *living_street*, and *pedestrian*). The lower $Dist_{topo}$ value means that the place has a high potential to reach higher pedestrian costumers. Finally, since Öztürk Hacar et al. [19] mentioned about lower effect of the direct geometric properties of the closest street, the sinuosity and street length were removed from the prediction model.

Primary goal is to predict visitor scores: rating, which are classified into two categories (i.e., High and Low) using quantile-based segmentation. We employed a random forest classifier, a robust model capable of handling heterogeneous features by using the interrelations among spatial characteristics, network properties, and synthetic-PC. For reproducibility, the random_state parameter is fixed at 42, while the model is configured with 100 trees (n_estimators = 100), the Gini impurity criterion for split evaluation, and the square root of the total number of features considered at each split [26]. Additionally, the dataset is partitioned into training and testing sets using an 80–20% split, with a fixed random seed of 42 to ensure consistent data splitting across experiments.

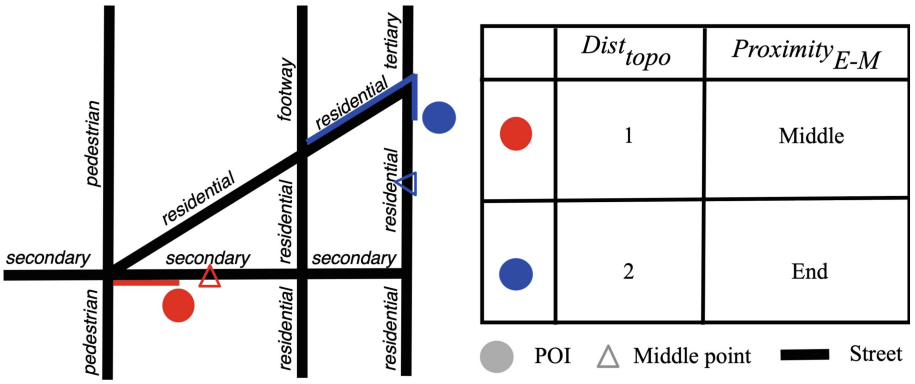

Fig. 3. $Proximity_{E\text{-}M}$ and $Dist_{topo}$ feature descriptions.

3 Experimental Results

Following the methodological framework outlined in Sect. 2, the proposed approach was applied to the Bakirköy study area to evaluate how various spatial features, synthetic-PC, and centrality measures relate to digital consumer behavior at EDO places. The procedures involved extracting and processing POI information from Google Maps, generating an arc-node topology for the OSM road network, and computing both distance-based and network-based features to serve as independent variables for the random forest classification model.

This study employs a selected set of centrality measures, each highlighting a distinct facet of network structure. *B* emphasizes those segments most frequently used in shortest paths, often corresponding to key thoroughfares. *C* identifies nodes that are situated relatively near all other nodes, while *D* highlights the number of direct connections at each node. *E* reflects a node's importance based on the influence of its neighbors, whereas *H* modifies the concept of closeness by placing greater emphasis on short distances. *K* further extends eigenvector centrality by introducing a damping factor that weighs distant connections. *L* is similar to betweenness but considers distributed flow rather than single shortest paths. Finally, *P* evaluates node importance by factoring in the significance of inbound connections [27–40]. Figure 4 illustrates the distribution of these eight centrality metrics across the street network. In each subfigure, streets are rendered in red, green, and blue to represent higher or lower centrality values. Taken together, as Öztürk Hacar et al. [19] stated, these distributions reveal how certain streets stand out as major connectors within Bakirköy's urban layout, remarking the network's structural hierarchy and the potential influence of these thoroughfares on consumer movement on the network.

Table 1 summarizes the prediction model's classification performance for the Low and High rating categories. The model is notably effective at identifying venues that fall into the Low rating category. However, the result on the prediction of the High rating category gives a more moderate performance in recognizing the places. Overall, the weighted F1-score of 0.76 demonstrates a balanced level of accuracy across both classes, although there is room for improvement in capturing High-rated venues.

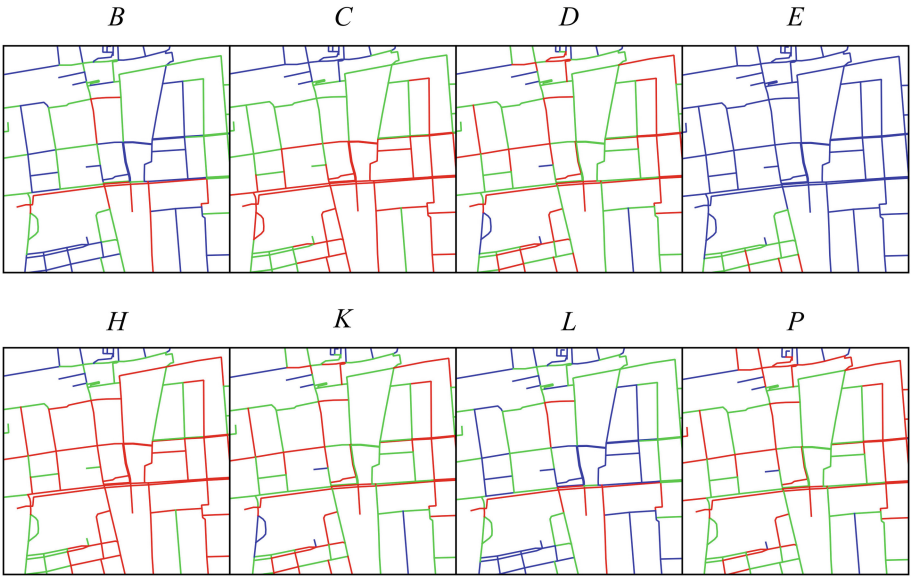

Fig. 4. Distribution of centrality measures across the street network: red as high, green as midrange, and blue as low.

Table 1. Performance metrics of the prediction model

Rating	Precision	Recall	F1-score
Low	0.79	0.86	0.83
High	0.70	0.58	0.64
Weighted	0.76	0.76	0.76

Figure 5 presents the mean decrease in impurity for each feature, revealing how they contribute to distinguishing between High and Low rating classes. $Dist_E$, the distance to a potentially densified junction, emerges as the most influential factor, implying that EDO places located nearer to such junctions are more likely to exhibit distinct rating patterns. Voronoi area and synthetic-PC follow closely, highlighting the relevance of local catchment size and approximate pedestrian activity for rating prediction. The categorical type of an EDO establishment also shows substantial importance, suggesting that different venue types attract varied consumer feedback. Centrality metrics: B, H, E, L, P, K, and C occupy mid-range positions, indicating that network connectivity and integration exert a measurable, though less dominant, effect. Meanwhile, $Proximity_{E\text{-}M}$, functional road classification, topological distance to pedestrian-friendly roads ($Dist_{topo}$), and the centrality metric D rank lower, but still contribute to the classification. Notably, the black lines on each bar represent the standard deviation, illustrating the variability of these importance scores across the ensemble. While $Dist_E$ and synthetic-PC exhibit relatively small deviations, indicating more consistent contributions, certain centrality

metrics display wider error lines, suggesting that their impact may fluctuate depending on the specific data subsets used in individual decision trees.

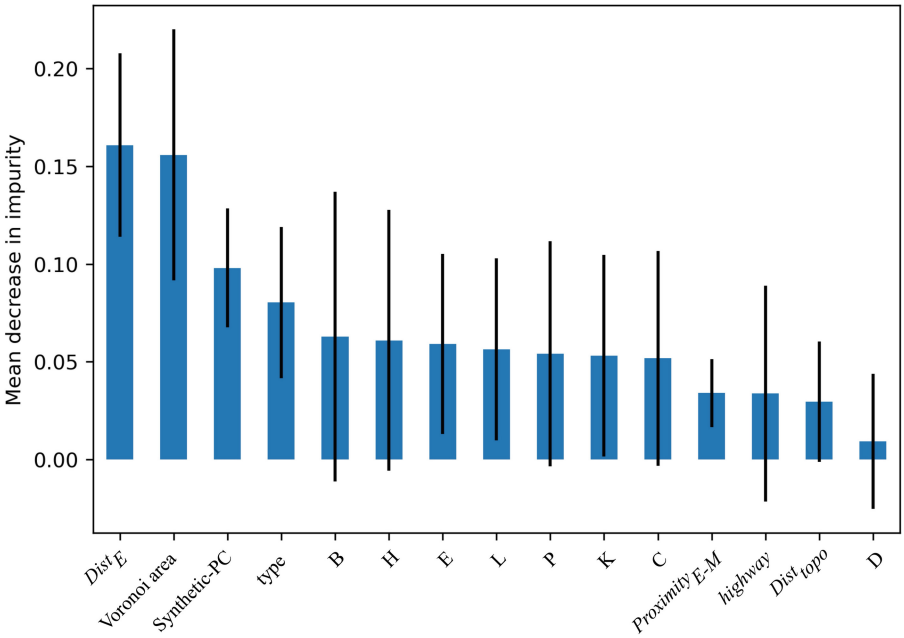

Fig. 5. Sorted feature importances with standard deviation across trees.

4 Conclusion

This study presents a novel approach for predicting visitor ratings at EDO venues by leveraging integrated spatial data derived from Google POIs and an OSM road network, along with a synthetic pedestrian measure. The approach was tested in Bakırköy, Istanbul, a district known for its high pedestrian flows due to the concentrations of educational institutions and shopping facilities. It offers key urban characteristics that influence digital consumer behavior. The analysis demonstrated that factors such as the proximity to densified junctions, local catchment areas, and synthetic pedestrian data play a pivotal role in shaping consumers' digital ratings, with centrality and various spatial metrics further complementing these insights.

The proposed approach can be employed as a decision support tool for entrepreneurs considering the location of new EDO business. By estimating the potential consumer rating at various candidate properties, entrepreneurs can identify the most promising sites based on projected visitor scores. This capability allows for data-driven decisions that align with consumer preferences, ultimately supporting higher success rates for newly established venues.

A key strength of this study is its avoidance of sensor-based, cellular network, or CCTV data collection, thereby reducing financial overheads and privacy concerns. Nonetheless, several limitations warrant attention. The synthetic pedestrian measure may not capture short-term time changes during a day in foot traffic. Future work should consider additional data inputs, such as temporal factors, land use variations, and consumer demographics, to validate and expand upon the present findings. Overall, this research offers a practical, privacy-conscious, and cost-effective strategy for understanding and forecasting digital consumer behavior, supporting the way for more informed urban planning and commercial decision-making.

Disclosure of Interests. The author declares no conflict of interest.

References

1. Valanarasu, M.R.: Comparative analysis for personality prediction by digital footprints in social media. J. Inform. Technol. Digit. World. **3**, 77–91 (2021)
2. Adikari, A., Nguyen, S., Nawaratne, R., De Silva, D., Alahakoon, D.: Transforming customer digital footprints into decision enablers in hospitality. Appl. Sci. **14**(3114) (2024). https://doi.org/10.3390/app14073114
3. Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., Van Alstyne, M.: Computational social science. Science. **323**, 721–723 (2009). https://doi.org/10.1126/science.1167742
4. Pentland, A.S.: The data-driven society. Sci. Am. 309, 78–83 (2013). https://www.jstor.org/stable/26018109
5. Di Clemente, R., Luengo-Oroz, M., Travizano, M., Xu, S., Vaitla, B., González, M.C.: Sequences of purchases in credit card data reveal lifestyles in urban populations. Nat. Commun. **9**, 3330 (2018). https://doi.org/10.1038/s41467-018-05690-8
6. Khan, A.M., Loan, F.A.: Exploring the reviews of Google Maps to assess the user opinions about public libraries. Libr. Manag. **43**, 601–615 (2022). https://doi.org/10.1108/LM-05-2022-0053
7. Hacar, M., Öztürk Hacar, Ö., Altafini, D., Pappalardo, L., Gülgen, F., Cutini, V.: From streets to screens: A deep dive into urban configuration and Google reviews. In: Proceedings of the 14th Space Syntax Conference, Nicosia, Cyprus (2024)
8. Muhammad, S.S., Dey, B.L., Weerakkody, V.: Analysis of factors that influence customers' willingness to leave big data digital footprints on social media: a systematic review of literature. Inf. Syst. Front. **20**, 559–576 (2018). https://doi.org/10.1007/s10796-017-9802-y
9. Girardin, F., Calabrese, F., Fiore, F., Ratti, C., Blat, J.: Digital footprinting: uncovering tourists with user-generated content. IEEE Pervasive Comput. **7**, 36–43 (2008). https://doi.org/10.1109/MPRV.2008.71
10. Martí, P., García-Mayor, C., Serrano-Estrada, L.: Taking the urban tourist activity pulse through digital footprints. Curr. Issue Tour. **24**, 157–176 (2021). https://doi.org/10.1080/13683500.2019.1706458
11. Jia, Y., Liu, I.: Do consumers always follow "useful" reviews? The interaction effect of review valence and review usefulness on consumers' purchase decisions. J. Assoc. Inf. Sci. Technol. **69**, 1304–1317 (2018). https://doi.org/10.1002/asi.24050

12. Li, H., Hecht, B.: 3 stars on yelp, 4 stars on Google Maps: a cross-platform examination of restaurant ratings. Proc. ACM Human Comput. Interact. **4**, 254 (2021). https://doi.org/10.1145/3432953
13. Akkaya, M., Özçevik, Ö., Tepe, E.: A machine learning application to Google Maps Reviews as a participatory planning tool. Int. J. Urban Sci. **28**, 379–402 (2024). https://doi.org/10.1080/12265934.2024.2320916
14. Owuor, I., Hochmair, H.H.: An overview of social media apps and their potential role in geospatial research. ISPRS Int. J. Geo Inf. **9**(526) (2020). https://doi.org/10.3390/ijgi9090526
15. Tang, Y., Zhang, J., Liu, R., Li, Y.: Exploring the impact of built environment attributes on social followings using social media data and deep learning. ISPRS Int. J. Geo Inf. **11**(325) (2022). https://doi.org/10.3390/ijgi11060325
16. Rajabi, F., Hosseinali, F., Rabiei-Dastjerdi, H.: An examination and analysis of the clustering of healthcare centers and their spatial accessibility in Tehran Metropolis: Insights from Google POI Data. Sustain. Cities Soc. **117**, 105845 (2024). https://doi.org/10.1016/j.scs.2024.105845
17. Tse, M.P., Dhalla, I., Nayyar, D.: Google star ratings of Canadian hospitals: a nationwide cross-sectional analysis. BMJ Open Qual. **13**, e002713 (2024). https://doi.org/10.1136/bmjoq-2023-002713
18. Leiras, A., Eusébio, C.: Perceived image of accessible tourism destinations: a data mining analysis of Google Maps reviews. Curr. Issue Tour. **27**, 2584–2602 (2024). https://doi.org/10.1080/13683500.2023.2230338
19. Öztürk Hacar, Ö., Hacar, M., Gülgen, F., Pappalardo, L.: Where we rate: the impact of urban characteristics on digital reviews and ratings. Appl. Sci. **15**(2), 931 (2025). https://doi.org/10.3390/app15020931
20. Jiang, F., Ma, J., Li, Z.: Pedestrian volume prediction with high spatiotemporal granularity in urban areas by the enhanced learning model. Sustain. Cities Soc. **79**, 103653 (2022). https://doi.org/10.1016/j.scs.2021.103653
21. Google LLC, Google Maps API, Available online: https://developers.google.com/maps last accessed 15 Mar 2025
22. OpenStreetMap, OSM Export, Available online: https://www.openstreetmap.org/export last accessed 15 Mar 2025
23. QGIS, Firenze Version 3.28 LTR, Available online: http://www.qgis.org/en/site/index.html last accessed 15 Mar 2025
24. OpenStreetMap Wiki, Map Features, Available online: https://wiki.openstreetmap.org/wiki/Map_Features last accessed 15 Mar 2025
25. NetworkX Python Package, Available online: https://networkx.org/ last accessed 15 Mar 2025
26. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: machine learning in Python. J. Mach. Learn. Res. **12**, 2825–2830 (2011)
27. Porta, S., Crucitti, P., Latora, V.: The network analysis of urban streets: a primal approach. Environ. Plan. B Plan. Design. **33**(5), 705–725 (2006). https://doi.org/10.1068/b32045
28. Buhl, J., Gautrais, J., Reeves, N., Solé, R.V., Valverde, S., Kuntz, P., Theraulaz, G.: Topological patterns in street networks of self-organized urban settlements. Eur. Phys. J. B. **49**, 513–522 (2006). https://doi.org/10.1140/epjb/e2006-00085-1
29. Crucitti, P., Latora, V., Porta, S.: Centrality measures in spatial networks of urban streets. Phys. Rev. E. **73**, 036125 (2006). https://doi.org/10.1103/PhysRevE.73.036125
30. Jiang, B.: A topological pattern of urban street networks: universality and peculiarity. Phys. A Statist. Mech. Appl. **384**(2), 647–655 (2007). https://doi.org/10.1016/j.physa.2007.05.064
31. Barthélemy, M., Flammini, A.: Modeling urban street patterns. Phys. Rev. Lett. **100**, 138702 (2008). https://doi.org/10.1103/PhysRevLett.100.138702

32. Masucci, A.P., Smith, D., Crooks, A., Batty, M.: Random planar graphs and the London street network. Eur. Phys. J. B. **71**, 259–271 (2009). https://doi.org/10.1140/epjb/e2009-00290-4
33. Masucci, A.P., Stanilov, K., Batty, M.: Limited urban growth: London's street network dynamics since the 18th century. PLoS One. **8**, e69469 (2010). https://doi.org/10.1371/journal.pone.0069469
34. Strano, E., Nicosia, V., Latora, V., Porta, S., Barthélemy, M.: Elementary processes governing the evolution of road networks. Sci. Rep. **2**(296) (2012). https://doi.org/10.1038/srep00296
35. Corcoran, P., Mooney, P., Bertolotto, M.: Analysing the growth of OpenStreetMap networks Spat. Statist. 3, 21–32 (2013) 10.1016/j.spasta.2013.01.002
36. Zhao, P., Jia, T., Qin, K., Shan, J., Jiao, C.: Statistical analysis on the evolution of OpenStreetMap road networks in Beijing. Physica A. **420**(59–72) (2015). https://doi.org/10.1016/j.physa.2014.10.076
37. Hacar, M., Kılıç, B., Şahbaz, K.: Analyzing OpenStreetMap road data and characterizing the behavior of contributors in Ankara, Turkey. ISPRS Int. J. Geo-Inform. **7**(10), 400 (2018). https://doi.org/10.3390/ijgi7100400
38. Freeman, L.C.: A set of measures of centrality based upon betweenness. Sociometry. **40**, 35–41 (1977)
39. Brandes, U.: On variants of shortest-path betweenness centrality and their generic computation. Soc. Networks. **30**, 136–145 (2008). https://doi.org/10.1016/j.socnet.2007.11.001
40. Hacar, M., Altafini, D., Cutini, V.: Network-based hierarchical feature augmentation for predicting road classes in OpenStreetMap. ISPRS Int. J. Geo Inf. **13**(12), 456 (2024). https://doi.org/10.3390/ijgi13120456