BCT Journal

Journal of Business, Communication & Technology

Collaborative Advertisement Recommendation System Leveraging User Preferences, Geography, and Demographics

Djalila Boughareb[1*], Hazem Bensalah[2], Zineddine Kouahla[1]
[1]University of 8 Mai 1945, Algeria, [2] University of El Oued, Algeria

Abstract The increasing demand for personalized advertising based on user preferences is driving a surge in popularity. Social networks utilize millions of user' data to suggest ads based on specific criteria. However, many of these ads can be uninteresting. This paper presents a collaborative advertisement recommendation system that leverages users' preferences along with geographic and demographic data to deliver engaging ads. The system employs the K-dtree algorithm to efficiently organize users into interest-based communities and model complex patterns within those communities to enhance ad relevance. The dataset, collected via Hazmit provides a rich source of information. The system's performance was evaluated based on precision, recall, F-score, and accuracy metrics, as well as running time measurements. The results highlighted the superior effectiveness of the K-dtree-based approach in accurately targeting the right customers for advertisements. Overall, the K-dtree method improves ad targeting accuracy, especially for food and demographics, but struggles with news due to subjectivity and regional biases.

Keywords: *Recommender systems, Advertisement recommendation, K-dtree, Collaborative filtering*

[*]Corresponding Author:
Djalila Boughareb
dboughareb@gmail.com

Received: February 2025
Revised: April 2025
Accepted: May 2025
Published: May 2025

1. Introduction

Recently, recommender systems have become increasingly important in recommender algorithms and methods. Recommender system can be described as programs that attempt to recommend the most suitable items (products or services) for unique users (individuals or companies) by anticipating the user's interest in the item based on data related to items, users, and interactions between items and users (Ahuja et al., 2019).

In recent years, the rapid rise in social network (SN) advertising has been driven by platforms' ability to recommend ads based on predefined criteria, such as geographic, demographic, and user preference data (Piatykop & Pronin, 2020). These systems rely on vast datasets collected from millions of users, yet many recommended ads remain uninteresting, repetitive, or boring—

http://doi.org/10.56632/bct.2025.4102

a challenge exacerbated by inefficiencies in traditional targeting models (Jibril & Adzovie, 2022). Advertisers aim to suggest adapted advertisements to increase sales and avoid financial waste caused by poorly targeted campaigns (Piatykop & Pronin, 2020). To address these gaps, this research proposes an advertisement recommendation system that synthesizes geographic-based, demographic-based, and user preference data, leveraging hierarchical algorithms like the K-d tree for granular segmentation (Ahuja et al., 2019; Das et al., 2019; Gong, 2010; Zarzour et al., 2018).

The K-d tree algorithm, known for its utility in multi-dimensional data organization, enables hierarchical clustering to explore user clusters at varying granularities. This is particularly valuable in collaborative recommendation systems, where users exhibit diverse degrees of similarity and preference (Ahuja et al., 2019). Unlike purely AI-driven models (e.g., LSTMs for ad optimization; Gummadi et al., 2024), this hybrid approach balances interpretability with scalability, addressing the “cold start” problem through demographic heuristics while aligning with the Uses and Gratifications Theory, which emphasizes user-centric media engagement (Jibril & Adzovie, 2022). By prioritizing multi-criteria data (geographic, demographic, and preference-driven), the system reduces redundancy and enhances novelty—critical for sustaining engagement in saturated digital ecosystems (Sun et al., 2025).

Furthermore, hierarchical clustering mitigates algorithmic monotony, a key contributor to user fatigue, by diversifying recommendations without sacrificing relevance (Jibril & Adzovie, 2022). This strategy aligns with ethical frameworks advocating transparency in data usage, a growing concern in AI-driven advertising (Gummadi et al., 2024). Advertisers stand to benefit from optimized targeting, reduced financial waste, and improved conversion rates, while platforms gain competitive advantages through enhanced user satisfaction (Piatykop & Pronin, 2020; Gummadi et al., 2024). In summary, the study seeks to explore the following research questions.

RQ1. How do the proposed centroid and hierarchical approaches compare in performance?

RQ2. What are the primary attributes of users, and which types of advertisements among the four daily interests (food, clothing, high technology, and news) are preferred by users based on geographic and demographic factors?

2. Theoretical Framework

Recommender systems try to understand relationships that provide referral systems with a good understanding of customers. There are three main types of relationships (Shetty, 2019): (1) User-product relationship which occurs when certain users have an affinity or preference for specific products that they need (Boughareb & Farah, 2012). For example, a soccer player might have a preference for soccer-related articles, so the social media site will build a player-to-soccer user-product relationship. (2) Product-product relationship, which occurs when items are of a similar nature, either in appearance or description. Some examples include books or music of the same genre, dishes from the same kitchen, or news articles about a particular event. (3) User-user relationship, which occurs when certain customers have similar tastes about a particular product or service, e.g., mutual friends, similar preferences, the same location, etc. Indeed, there are several techniques of recommendation, some of which are explained below.

(i) Content-Based Recommendation (CB): In this technique, items that are similar to items preferred by a specific user are recommended. Two methods are used to generate recommendations. The first is simple, whereby measures of similarity are used, and the other method generates recommendations using machine or deep learning methods, which are largely constructive models that are able to learn a user’s interests by having a history of their past interests (Lu et al., 2015).

In the work of Shu et al. (2018), the authors proposed a CB recommendation algorithm based on convolutional neural networks (CNN) to allow students to discover new learning resources that match their tastes and enable the e-learning system to recommend the learning resources to the right students. Moreover, Suglia et al. (2017) investigated the effectiveness of Recurrent Neural Networks (RNNs) in a top-N content-based recommendation. More recently, Kim et al. (2021) proposed an advertising video recommendation procedure based on computer vision and deep learning. The different changes in users’ facial expressions are captured at every moment and used to represent a user's emotions toward advertisements. For such a reason, a CNN-based prediction model to predict ratings, as well as a SIFT algorithm-based similarity model, was developed. The results show the utility of facial expression in properly addressing the new user problem in existing recommender systems. More recently, Boughareb et al. (2023b) investigated the efficacy of leveraging large Knowledge Graphs (KGs) for interpretable recommendations through Graph Attention Networks (GATs). Authors tried to maximize the utilization of semantic information and retain inherent knowledge encoded in relations by jointly learning embeddings for nodes (entities) and edges (properties). By integrating original data with additional knowledge from the Linked Open Data (LOD) cloud and applying GATs to generate node representations, they achieved promising results in the top-K recommendation task across three real-world datasets. The proposed approach capitalized on the rich structured information within KGs to provide explanations for recommendations.

(ii) Collaborative Filtering-Based Recommendation (CF): In this technique, recommended items are chosen based on the ratings and preferences of other people who share similar interests like country, sex, age, or like the same items. Gong (2010) proposed a personalized recommendation approach based on the k-means algorithm. The main idea consists of combining user clustering and item clustering technologies. Users are grouped based on their ratings of items, and each user group has a cluster center. Based on the similarity between the target user and the cluster centers, the target user's nearest neighbors can be identified and the predicted smoothed as needed. The suggested method used item clustering collaborative filtering to generate suggestions. The results of the experiments performed on the MovieLens dataset show that the collaborative filtering recommendation combining user clustering and item clustering is more scalable and accurate than the old one.

De Maio et al. (2020) proposed a context-aware advertising recommendation model on Twitter for determining whether a user u at the time t and in a specific location m might be interested or not in a given advertisement. They used the social network platform to expose some products to potentially interested users. The system uses text analysis services to extract knowledge from social network posts and advertisements as well as a location-based system. Also, Jouyandeh and Zadeh (2022) focused on advertising images. They used social media analysis to identify and extract the features out of an image in order of importance and try to find its potential customers who might be interested in by analyzing the content shared by users. The cold start problem is one of the main problems of collaborative filtering. It occurs when the user community is small and the data set is sparse (Huang et al., 2004). The great development of social media contributes to providing more information about the user, which helps to improve CF-based recommendations. In fact, the logic and the reality in which we live tell us that most people buy products suggested by those who are close to them.

Logesh et al. (2018) proposed to cluster the users of the Yelp and the TripAdvisor dataset by using bio-inspired techniques individually and producing the final clustering results using a statistical ensemble model. Then, a neighborhood search was performed with respect to the active target user in order to include it in the highly matched cluster. The ratings were then approximated based on the active target user's current neighbors in the cluster. The top-n list of suggestions was constructed and displayed to the user. Zarzour et al. (2018) proposed to create users' clusters using k-means; then for each obtained cluster, they applied Singular Value Decomposition (SVD) for Dimensionality Reduction to obtain decomposition matrices and calculated the similarity between matrices to get the final the recommendation model.

Ahuja et al. (2019) suggested the employment of the Clustered Sum of the Squared (CSS) method to determine the appropriate number of clusters, and then using k-means, they constructed a utility clustered matrix, which was used to compute the Pearson correlation similarity between users. Finally, they applied KNN to predict movies for the target user using the utility clustered matrix. Das et al. (2019) proposed a Collaborative Filtering based recommender system, which used a k-dtree for partitioning the users' space with respect to the location. Experiments conducted on the MovieLens dataset showed the effectiveness of the approach with a considerable reduction in the running time. More recently, Boughareb et al. (2023a) introduced FindLoc, a collaborative recommendation system tailored to assist Algerian users in discovering appealing destinations. Utilizing user evaluations, including ratings and comments, gathered from diverse locations within a city, the authors employed sentiment analysis to discern positive and negative sentiments from the collected users' interactions. The recommendation algorithm leveraged location data, previous ratings, user feedback, and age information to suggest intriguing places (doctor, restaurant, school, dentist, cafeteria, etc). Evaluation by real users yielded promising results, with a precision rate of 0.91, showcasing the system's effectiveness in facilitating personalized location recommendations.

(iii)Hybrid Recommendation: In order to achieve higher performance and overcome the limits of traditional recommendation techniques, hybrid approaches are usually proposed. They combine both techniques: content-based filtering approaches and collaborative filtering approaches. Such a hybrid approach may help to surmount the limitations of both approaches. In fact, the CF approach is not capable of providing recommendations in cold-start situations (Schein et al., 2002). In an attempt to avoid cold-start, sparseness, and/or scalability problems, hybrid approaches try to combine the best characteristics of two or more recommendation techniques into one hybrid technique (Adomavicius & Tuzhilin, 2005; Burke, 2002).

According to Burke (2007), there are seven hybridization mechanisms used in recommender systems to build hybrids: weighted (Mobasher et al., 2004), mixed (Smyth & Cotter, 2000), switching (Billsus & Pazzani, 2000), feature combination, feature augmentation (O'Sullivan & Smyth, 2004; Wilson & smyth, 2003'), cascade (Pazzani, 1999) and meta-level (Bellogin et al., 2013). Advertisements appear on many sites and are a great source of profit for them. Many companies pay huge sums of money to these sites to publish their ads on them, and they pay more if their ads get more clicks from users. Many methods do not care about the opportunity for the user to benefit from the displayed content. They only display it. In their work, Liu et al. (2018) focused on banner ads that are generally placed on news sites and took into account the user's satisfaction with the ad that appears to him, using recommendations and reasonable fairness for the frequency of appearances for each advertisement. Furthermore, in order to assign advertisements to groups that represent different media channels, the semantic and sentiment content of advertisements were modeled using a neural network classifier (Abrahams et al., 2013; Liu et al., 2018). More recently, Asfar et al. (2022), Chen et al. (2023) and Deldjoo et al. (2024) have conducted an in-depth analysis of contemporary recommendation algorithms and their applicability, furnishing a comprehensive survey on their recent developments and relevance in diverse contexts.

To the authors' best knowledge, there are no studies on advertisement recommendations that combine age, preferences, and position to classify them. In addition, different positions (east, west, north, and south) were not used before as a classification feature. Even by utilizing similarities within hierarchical clusters, hierarchical-based techniques can aid in reducing sparsity problems. Actually, dealing with data sparsity may be quite difficult, especially when it comes to long-tail products or customer preferences.

3. Methodology

3.1. Dataset

Collecting datasets from social media platforms has challenges regarding privacy and data availability, with some platforms restricting historical data access. To overcome this, a dataset was acquired from Hazmit, a social network tailored for the study's purposes. Hazmit allows users to create profiles and post various types of content, with features such as like/dislike buttons and comments for evaluation. To safeguard privacy, numerical identifiers were assigned to user profiles. The dataset comprised 498 interactions from 171 unique users, covering likes, dislikes, and ratings. Each user descriptor included features such as age, gender, city position, and ad ratings. Advertisements included videos and posters in domains like technology, food, clothing, and news, reflecting significant aspects of consumers' lives and interests, making them effective choices for advertising strategies. Table 1 exemplifies a data descriptor, showcasing the structure and content of the dataset.

Table 1

A Sample of the Dataset

UserId	Age	Sex	East	North	West	Technology	Food	Clothes	News
1	22	1	1	1	0	4	3	5	1
2	21	1	1	1	0	4	3	5	1
3	25	1	1	1	0	4	3	5	1
						...		...	
171	51	0	1	1	1	2	3	5	5

The dataset includes 3 binary features: sex with values 1 representing male and 0 representing female, age with values 1 including users higher than 30 years old and 0 including users lower than 30 years, and position with values 1 indicating east/north and 0 indicating west/south. Additionally, it contains ratings for advertisements across four domains: technology, food, clothing, and news.

Figure 1

The Distribution of Users Based on Their Gender, Age, and Geographical Location

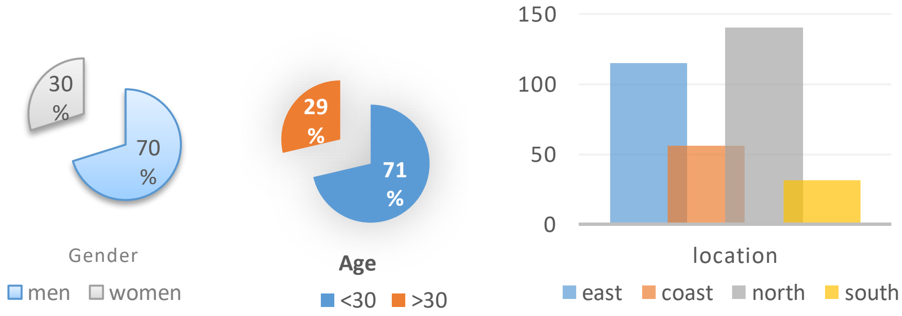


Figure 1 illustrates the distribution of users based on gender, age, and geographical location. The dataset comprises 120 men and 51 women. Geographically, users are spread across regions: 115 in the east, 56 in the west, 140 in the north, and 31 in the south. Users' ages range from 17 to 65 years old; 49 users are older than 30 years old, and 122 are younger than 30 years old.

3.2. Communities' Creation and Ads Recommendation

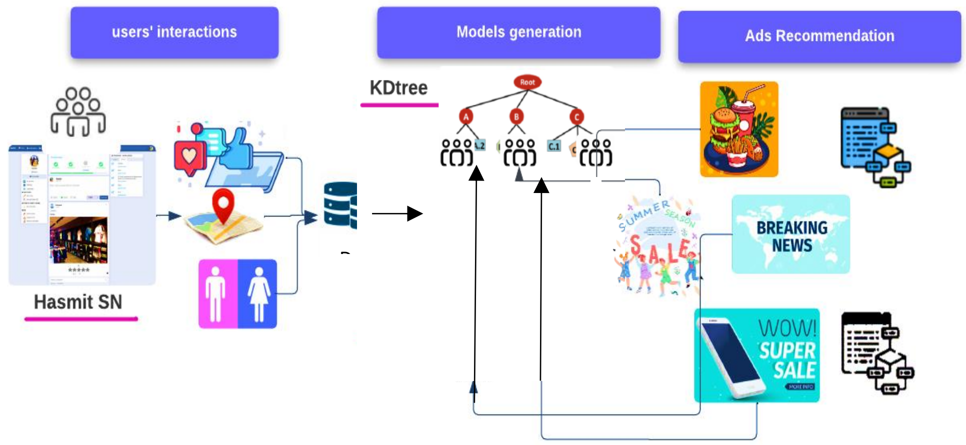
The ad recommender system begins with user registration, where users provide demographic information such as gender, age, and location and engage with ads through ratings, likes, and comments. This interaction data is collected and processed. For user classification, the K-dtree method is employed. The K-dtree algorithm groups users with similar traits, allowing for efficient neighbor searches within the demographic space. Based on these groupings, the system suggests ads tailored to user profiles and interests. As users interact with the recommended ads, their feedback helps refine the system's accuracy. Figure 2 illustrates the role of the K-dtree algorithm in user classification and community identification within the system.

3.3. K-dtree and K-dtree-Based Recommendation

A binary tree with k nodes, each of which represents a k-dimensional point, is called a k-dimensional (k-d). In both 2D planes and 3D spaces, it is frequently used to divide collections of objects. At the beginning, two areas, "b" and "c," are created from the entire space "a" (level 0), according to x-coordinate values. Following that, y-coordinate values are used to further divide each of these sections into two. Using different partitioning criteria each time, this procedure is repeated until the tree is completely built. Applications for K-dtrees can be found in many fields.

Although, the K-dimensional tree (Shu et al., 2018) is a binary tree where each node is a K-dimensional point. It is possible to consider any non-leaf node as inherently generating a splitting hyperplane that splits the space into two equal halves or half-spaces. The left subtree of that node is represented by points to the left of this hyperplane, and the right subtree is represented by points to the right of this hyperplane. It can be used to partition a set of objects into a 2D plane and a 3D space.

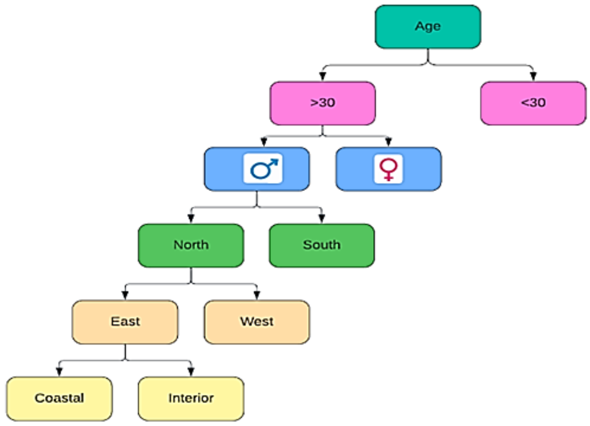
Figure 2
The System's Workflow



The effectiveness of K-dtrees in the domain of recommender systems has been proven (Das et al., 2013; Das et al., 2019). For example, Das et al. (2013) proposed a location-based recommender algorithm that partitions users' data based on a K-dtree data structure. Also, the work of Das et al. (2019) consists of using K-dtree and Quadtree to partition the set of users into smaller clusters based on location and apply CF-based recommender algorithms individually to these clusters. Both of these works have obtained very good results.

The K-dtree-based recommendation system introduces a unique approach to recommendation generation by utilizing a hierarchical, demographic-driven K-dtree structure. The originality of this system lies in how it partitions user data into communities based on age, sex, and geographical position, which is uncommon in traditional recommendation algorithms. Initially, the dataset is split into training and test sets (80-20 ratio), after which a K-dtree is constructed, with each node representing a community of users. At the first level, users are grouped into two broad categories based on age (above and below the mean age). At the second level, these age groups are further subdivided by sex, resulting in four distinct communities (younger men, older men, younger women, older women). The third level introduces geographical divisions (e.g., northeast position), refining the communities into eight distinct subgroups. Figure 3 illustrates the tree structure.

Figure 3
The K-dtree Structure



This structured approach enables efficient user segmentation, with the K-dtree's logarithmic search complexity O(log n) drastically reducing the time required to find the nearest demographic matches. The key innovation of the K-dtree-Rec algorithm lies in how it leverages this structure to identify the K closest users based on demographic similarities (age, sex, and position) and uses their aggregated preferences to make recommendations. By recursively dividing users based on demographic data, the system generates highly personalized recommendations, specifically targeting users who share similar characteristics. The algorithm calculates average ratings for advertisements or products by aggregating ratings from similar users, offering tailored predictions that reflect collective preferences.

Algorithm: k-dtree-Rec
Input: city, age, sex of the user
Output: recommendation
1. Perform a search in the k-dtree structure to identify the n closest users to the current user based on demographic attributes.
2. Calculate the average rating for each advertisement by aggregating the ratings provided by the n closest users.
3. Present the predicted ratings for the advertisements.
End.

After users are classified using the k-dtree algorithm, a selection of advertisements is tailored and proposed to them. Ad topics cover four categories: food, technology, clothing, and news. Each user receives personalized ad recommendations based on average ratings from similar peers within their community. For instance, if nearby users give ratings of 3, 5, and 2 to technology ads, the projected rating for technology ads will be 3.33. Table 2 displays the computed ratings for each ad domain. For example, user 9 receives strong endorsements for clothing ads due to high ratings from similar users in the community.

Table 2
The Computed Rating of Each Domain

Id	Technology	Food	Clothes	News
9	3.266667	3.466667	3.666667	1.600000
21	4.284444	3.764444	4.244444	2.306667
25	3.885630	3.584296	4.082963	1.553778
….				
163	4.057188	3.777266	3.341446	2.864942
166	3.937146	3.451818	3.422763	2.790996

4. Results

4.1. Centroid versus Hierarchic Approach Performance Comparison

To address RQ1, performance metrics such as accuracy, precision, recall, and F-score are selected as crucial measurements, particularly in the context of ad recommendation, to evaluate the proposed approaches. The degree to which the suggested advertisements correspond with the user's tastes is reflected in the accuracy of the algorithm, which measures its overall correctness. Precision emphasizes the significance of avoiding inundating users with irrelevant adverts by measuring the percentage of appropriately recommended ads out of all recommended ads. In this instance, recall refers to the algorithm's capacity to pick up on all relevant adverts, making sure that recommendations that could be helpful aren't overlooked. By combining recall and precision, the F-score offers a fair evaluation of the algorithm's capacity to maximize pertinent recommendations while minimizing irrelevant ones. The evaluation metrics are represented by Eq. (2-5).

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$\text{Precision}=\frac{TP}{TP+FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP+FN} \qquad (4)$$

$$\text{F1-score} = \frac{TP}{TP+\frac{1}{2}(FP+FN)} \qquad (5)$$

Such as:

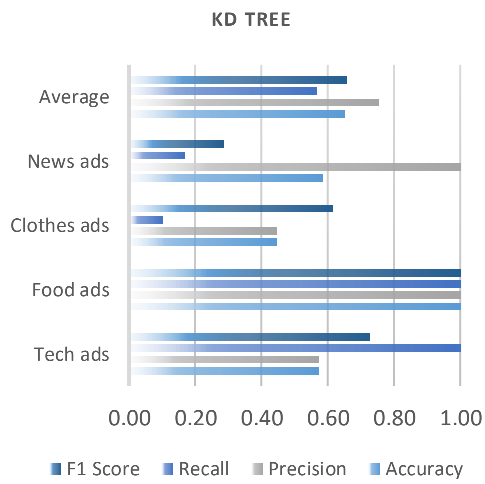
TP: True Positives (correctly identified positive cases)
FP: False Positives (incorrectly identified positive cases)
TN: True Negatives (correctly identified negative cases)
FN: False Negatives (incorrectly identified negative cases)

The results summarized in Table 3 and Figure 4 respectively show the accuracy, precision, recall, and F-score values according to the category.

Table 3
Prediction Performance in Terms of Accuracy, Precision, Recall, and F-Score

Categories	**K-dtree**			
	Accuracy	Precision	Recall	F1 Score
Tech ads	0.57	0.57	1.00	0.73
Food ads	1.00	1.00	1.00	1.00
Clothes ads	0.44	0.44	0.10	0.62
News ads	0.58	1.00	0.17	0.29
Average	0.65	0.75	0.57	0.66

Figure 4
Prediction Performance in Terms of Accuracy, Precision, Recall, and F-score



In the context of ad suggestion, the K-dtree's higher precision plays a crucial role in improving user engagement and satisfaction, particularly when reducing the number of irrelevant ads. By increasing precision, the system enhances the user experience and potentially raises click-through rates. The accuracy values obtained from the K-dtree algorithm provide insight into how well it performs in recommending various types of ads. For instance, food-related ads achieve a perfect accuracy score of 1, indicating that the K-dtree is highly effective at matching users with relevant food advertisements. However, K-dtree shows a lower accuracy of 0.44 for clothing ads, suggesting a noticeable gap between the recommended ads and users' actual preferences in this category. Tech and news ads receive accuracy scores of 0.57 and 0.58, respectively, reflecting a moderate alignment with user interests. These results demonstrate the varying levels of performance of the K-dtree algorithm across different ad types, with high precision in some areas and room for improvement in others.

The performance metrics for food ads using the K-dtree algorithm reveal its exceptional ability to deliver relevant recommendations. The K-dtree achieves perfect scores across all evaluation metrics: accuracy, precision, recall, and F-score, all at 1.00. This indicates that the K-dtree effectively recommends food ads that align perfectly with users' preferences, with no false positives or negatives. This flawless performance underscores the algorithm's effectiveness in accurately matching food ads to user interests.

However, the K-dtree faces challenges when recommending news ads, reflected in its lower accuracy and F-score values. These challenges can be attributed to the inherent subjectivity and complexity of news content, making it difficult for the algorithm to precisely forecast individual user preferences and engagement levels. The diversity and temporal sensitivity of news material complicates the process of capturing user interests effectively, resulting in reduced performance metrics in this category. Overall, while the K-dtree excels in recommending food ads, it encounters difficulties with news advertisements due to the dynamic and varied nature of news consumption.

Food advertisements achieved the highest accuracy using the K-dtree algorithm, demonstrating its effectiveness in aligning recommendations with user preferences in this category. However, news advertisements consistently received the lowest accuracy scores, indicating significant challenges in effectively recommending this type of content. This reduced performance may be attributed to a growing skepticism toward news-related marketing, as the prevalence of false information has led to a decline in public trust in news sources. Consequently, users may become less responsive to ads in the news category as they prioritize seeking out more reliable and trustworthy information. This decline in engagement underscores the importance of developing strategies to improve the relevance and credibility of news advertisements to regain user interest.

4.2. Impact of Geographic and Demographic Factors on User Preferences

4.2.1. Geographic Location

The results summarized in Table 4 show the accuracy, precision, recall, and F-score values based on the geographic location.

Table 4

Comparative Evaluation of the System Performance Based on Geographic Location

Region	Category	Accuracy	Precision	Recall	F1 Score
North	Tech	0.964	0.964	1	0.982
	Food	0.821	0.821	1	0.902
	Clothes	0.821	0.821	1	0.902
	News	0	0	0	0
South	Tech	0	0	0	0
	Food	0.833	0.833	1	0.909
	Clothes	0.333	0.333	1	0.5
	News	0	0	0	0
East	Tech	0.773	0.773	1	0.872
	Food	0.727	0.727	1	0.842
	Clothes	0.636	0.636	1	0.778
	News	0	0	0	0
Coast	Tech	0.833	0.833	1	0.909
	Food	1	1	1	1
	Clothes	0.917	0.917	1	0.957
	News	0	0	0	0

Regarding specific categories, all regions show strong accuracy, precision, recall, and F1 scores for the Tech and Food categories. Although somewhat less than in tech and food, clothes also fare well. The news category performance varies by region, with lower scores found in the East and South. Also, Performance measurements show significant disparities between regions, especially in the News category. In comparison to other regions, users from the South exhibit reduced interest in the tech and clothes categories. Compared to users in other regions, those from the East region show less interest in the news category.

4.2.2. Age

The results summarized in Table 5 show the accuracy values based on age.

Table 5
Evaluation Based on the Age

Accuracy	**Age >= 30**				**Age <30**			
	Technology	Food	Clothes	News	Technology	Food	Clothes	News
k-dtree	0.714	0.714	0.5	0.214	0.85	0.9	0.9	0

The results indicate that users under and over 30 exhibit distinct preferences for different types of advertisements. For users aged 30 and older, the K-dtree algorithm demonstrates notable effectiveness in promoting tech-related and food-related advertisements, achieving high to moderate accuracy scores—specifically, a 90.9% accuracy for tech ads and 81.8% accuracy for food ads. However, the algorithm's performance declines when recommending clothing ads, achieving only 50% accuracy in this category.

In contrast, younger users under 30 show a strong preference for food-related advertisements, with the K-dtree achieving a high accuracy of 90% in this area. While the K-dtree performs well for food ads among younger users, it struggles to effectively suggest news-related advertisements to this demographic. These findings underscore the necessity of tailoring ad recommendations to align with the unique preferences and demographics of different user groups. Understanding these variations is crucial for enhancing user engagement and ensuring that advertisements resonate with their intended audiences.

4.2.3. Gender

The results summarized in Table 6 show the accuracy, precision, recall, and F-score values segmented by gender (man or woman).

Table 6
Comparative Evaluation of K-Means and K-dtree Approaches Based on the Gender

Gender	**Category**	**Metrics**			
		Accuracy	***Precision***	***Recall***	***F1 Score***
Man	Technology	0.739	0.739	1.000	0.850
	Food	0.739	0.739	1.000	0.850
	Clothes	0.696	0.696	1.000	0.821
	News	0.378	0.250	0.490	0.330
Woman	Technology	0.909	0.909	1.000	0.952
	Food	1.000	1.000	1.000	1.000
	Clothes	0.818	0.818	1.000	0.900
	News	0.278	0.150	0.440	0.350

As Table 6 shows, the K-dtree algorithm demonstrates competency in recommending items across various categories. Specifically, it performs relatively better in the food category for women. The algorithm maintains competitive performance across all categories, with some variations in accuracy observed across different recommendation contexts.

Table 3 indicates that the K-dtree achieved higher accuracy overall. The gender-based analysis presented in Table 6 reveals that both men and women exhibit notable interest in high-tech, food, and clothing categories while displaying minimal interest in news advertisements. Furthermore, the presence of accuracy values reaching 100% in certain instances suggests that the impact of advertisements on women may be more significant compared to men.

We conclude by noting that users with similar traits, such as gender or geographic location, could be interested in clothing advertisements in various ways related to their preferences. People tend to mimic the behavior of those in their community regarding food and tech marketing, potentially due to societal influences on opinions. Additionally, the running time for the K-dtree is acceptable, at 1.10 seconds, while maintaining quality in recommendations.

5. Discussion

The rapid rise of SN advertising underscores the need for systems that balance personalization with user engagement (Piatykop & Pronin, 2020). Our study demonstrates that the K-dtree-based hierarchical approach achieves this balance, outperforming traditional centroid methods in accuracy (0.65 average) and precision (0.75 average) while maintaining computational efficiency (1.10 seconds runtime). These results align with the principles of the Uses and Gratifications Theory (Jibril & Adzovie, 2022), which emphasizes user-centric media engagement. By prioritizing geographic, demographic, and preference-driven data, our system reduces redundancy and enhances novelty, critical for sustaining engagement in saturated digital ecosystems (Sun et al., 2025).

The K-dtree's exceptional performance in food ads (100% accuracy, precision, recall, F1-score) reflects the algorithm's ability to leverage structured preferences (e.g., regional cuisines, dietary habits) and demographic homogeneity in this category. Conversely, the news ads' poor performance (0.29 F1-score) highlights challenges tied to subjectivity, temporal sensitivity, and declining user trust in news sources—a trend exacerbated by misinformation (Jibril & Adzovie, 2022). This disparity mirrors findings in AI-driven recommendation systems, where contextual complexity (e.g., news variability) often limits model efficacy compared to structured domains like food (Gummadi et al., 2024).

Our results reveal stark regional disparities: coastal regions achieved near-perfect accuracy for food (1.00) and clothes (0.917), likely due to homogenous consumer behaviors and cultural trends, while southern regions struggled with tech ads (0% accuracy), possibly reflecting infrastructural or socioeconomic divides. These geographic patterns align with studies on e-WoM and community-driven preferences, where localized social norms amplify engagement with culturally resonant ads (Jibril & Adzovie, 2022).

Demographic segmentation further validated the need for granular targeting. In terms of age, younger users (<30) exhibited stronger engagement with food ads (90% accuracy), while older users (≥30) favored tech (90.9% accuracy), reflecting generational divides in consumption priorities. In terms of gender, women showed higher responsiveness to food (100% accuracy) and clothes (0.818 accuracy) ads, likely due to culturally influenced purchasing roles, whereas men displayed moderate engagement across categories. These findings underscore the importance of hierarchical clustering in collaborative systems, where multi-granular exploration (Ahuja et al., 2019) mitigates algorithmic monotony and addresses the "cold start" problem through demographic heuristics.

While the K-dtree's interpretability and scalability are strengths, its limitations in news recommendations highlight ethical challenges tied to transparency and bias (Gummadi et al., 2024). For instance, the algorithm's inability to adapt to news skepticism (0% accuracy in some regions) risks amplifying user disengagement. Advertisers must balance hyper-personalization with ethical safeguards, such as diversifying training data and integrating user feedback loops to rebuild trust (Sun et al. 2025).

Disclosure Statement

The authors claim no conflict of interest.

Funding

The research did not receive any specific grants from funding agencies.

References

Abrahams, A. S., Coupey, E., Zhong, E. X., Barkhi, R., & Manasantivongs, P. S. (2013). Audience targeting by B-to-B advertisement classification: A neural network approach. *Expert Systems with Applications*, *40*(8), 2777–2791. https://doi.org/10.1016/j.eswa.2012.10.068

Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, *17*(6), 734–749. https://doi.org/10.1109/TKDE.2005.99

Ahuja, A., Solanki, A., & Nayyar, A. (2019). Movie recommender system using K-means clustering and K-nearest neighbor. In A. Bansal (Ed.), *9th International Conference on Cloud Computing, Data Science & Engineering* (pp. 263–268). IEEE. https://doi.org/10.1109/CONFLUENCE.2019.8776969

Asfar, M. M., Crump, T., & Far, B. (2022). Reinforcement learning-based recommender systems: A survey. *ACM Computing Surveys*, *55*(7), 1–38. https://doi.org/10.1145/3543846

Bellogin, I., Cantador, F., Díez, P., & Chavarriaga, E. (2013). An empirical comparison of social, collaborative filtering, and hybrid recommenders. *ACM Transactions on Intelligent Systems and Technology*, *4*(1), 1–29. https://doi.org/10.1145/2414425.2414439

Billsus, D., & Pazzani, M. (2000). User modeling for adaptive news access. *User Modeling and User-Adapted Interaction*, *10*, 147–180. https://doi.org/10.1023/A:1026501525781

Boughareb, D., & Farah, N. (2012). Contextual modelling of the user browsing behaviour to identify the user's information need. In E. Said, L. Jalal, & E. Ariwa (Eds.), *Proceedings of the Second International Conference on Innovative Computing Technology* (pp. 247–252). IEEE. https://doi.org/10.1109/INTECH.2012.6457773

Boughareb, D., Boughareb, R., Guerri, N., & Seridi, H. (2023a). FindLoc: A sentiment analysis-based recommender system. In U. Albalawi & S. A. Albelwi (Eds.), *3rd International Conference on Computing and Information Technology* (pp. 101–105). IEEE. https://doi.org/10.1109/ICCIT58132.2023.10273931

Boughareb, R., Seridi-Bouchelaghem, H., & Beldjoudi, S. (2023b). Joint representation of entities and relations via graph attention networks for explainable recommendations. *Journal of Web Engineering*, *22*(4), 615–638. https://doi.org/10.13052/jwe1540-9589.2243

Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction*, *12*, 331–370. https://doi.org/10.1023/A:1021240730564

Burke, R. (2007). Hybrid web recommender systems. In P. Brusilovsky, A. Kobsa, & W. Nejdl (Eds.), *The adaptive web* (pp. 377–408). Springer-Verlag. https://doi.org/10.1007/978-3-540-72079-9_12

Chen, X., Yao, L., McAuley, J., Zhou, G., & Wang, X. (2023). Deep reinforcement learning in recommender systems: A survey and new perspectives. *Knowledge-based systems*, *264*, Article 110335. https://doi.org/10.1016/j.knosys.2023.110335

Das, J., Majumder, S., & Gupta, P. (2013). Spatially aware recommendations using k-dtrees. In S. C. Satapathy & P. S. Avadhani (Eds.), *Proceedings of the 3rd International Conference on Computational Intelligence and Information Technology* (pp. 209–217). Institution of Engineering and Technology. https://doi.org/10.1049/cp.2013.2593

Das, J., Majumder, S., Gupta, P., & Mali, K. (2019). Collaborative recommendations using hierarchical clustering based on k-d trees and quadtrees. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *27*(4), 637–668. https://doi.org/10.1142/S0218488519500284

De Maio, C., Gallo, M., Hao, F., & Yang, E. (2020). Who and where: Context-aware advertisement recommendation on Twitter. *Soft Computing*, *25*, 379–387. https://doi.org/10.1007/s00500-020-05147-5

Deldjoo, Y., Schedl, M., & Knees, P. (2024). Content-driven music recommendation: Evolution, state of the art, and challenges. *Computer Science Review*, *51*, Article 100618. https://doi.org/10.1016/j.cosrev.2024.100618

Gummadi, V., Ramadevi, N., Udayaraju, P., Ravulu, C., Seelam, D. R., & Swamy, S. V. (2024). A deep learning-based optimization model for advertisement campaign. In J. Yogapriya (Ed.), *Proceedings of the 2024 5th International Conference on Smart Electronics and Communication (ICOSEC)* (pp. 1783–1790). IEEE. https://doi.org/10.1109/ICOSEC61587.2024.10722166

Gong, S. (2010). A collaborative filtering recommendation algorithm based on user clustering and item clustering. *Journal of Software*, *5*(7), 745–752.

Huang, Z., Chung, W., & Chen, H. (2004). A graph model for E-commerce recommender systems. *The Journal of the Association for Information Science and Technology*, *55*(3), 259-274. https://doi.org/10.1002/asi.10372

Jibril, A. B., & Adzovie, D. E. (2022). Understanding the moderating role of E-WoM and traditional media advertisement toward fast-food joint selection: A uses and gratifications theory. *Journal of Foodservice Business Research*, *27*(1), 1–25. https://doi.org/10.1080/15378020.2022.2070450

Jouyandeh, F., & Zadeh, P. M. (2022). IPARS: An image-based personalized advertisement recommendation system on social networks. *Procedia Computer Science*, *201*, 375–382. https://doi.org/10.1016/j.procs.2022.03.050

Kim, G., Choi, I., Li, Q., & Kim, J. (2021). A CNN-based advertisement recommendation through real-time user face recognition. *Applied Sciences*, *11*(20), Article 9705. https://doi.org/10.3390/app11209705

Liu, D. R., Liao, Y. S., Kuan, Y. H. C., & Chen, Y. (2018). Advertisement recommendation based on personal interests and ad push fairness. *Kybernetes*, *48*(8), 1586–1605.

Logesh, R., Subramaniyaswamy, V., Malathi, D., Sivaramakrishnan, N., & Vijayakumar, V. (2018). Enhancing recommendation stability of collaborative filtering recommender system through bio-inspired clustering ensemble method. *Neural Computing and Applications*, *32*, 2141–2164. https://doi.org/10.1007/s00521-018-3891-5

Lu, J., Wu, D., Mao, M., Wang, W., & Zhang, G. (2015). Recommender system application developments: A survey. *Decision Support Systems*, *74*, 12-32. https://doi.org/10.1016/j.dss.2015.03.008

Mobasher, B., Jin, X., & Zhou, Y. (2004). Semantically enhanced collaborative filtering on the web. In B. Berendt, A. Hotho, D. Mladenic, M. Someren, M. Spiliopoulou, & G. Stumme (Eds.), *Web mining: From web to semantic web* (pp. 57–76). Springer. https://doi.org/10.1007/978-3-540-30123-3_4

O'Sullivan, D., Smyth, B., & Wilson, D. (2004). Preserving recommender accuracy and diversity in sparse datasets. *International Journal on Artificial Intelligence Tools*, *13*, 219–235. https://doi.org/10.1142/S0218213004001491

Pazzani, M. (1999). A framework for collaborative, content-based and demographic filtering. *Artificial Intelligence Review*, *13*, 393–408. https://doi.org/10.1023/A:1006544522159

Piatykop, O., & Pronin, O. (2020). Model selection of the target audience in social networks in order to promote the product. *CEUR Workshop Proceedings*, *2604*, 1–10. https://ceur-ws.org/Vol-2604/paper29.pdf

Schein, A. I., Popescul, A., Ungar, L. H., & Pennock, D. M. (2002). Methods and metrics for cold-start recommendations. In K. Järvelin (Eds.), *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval'* (pp. 253–260). Association for Computing Machinery. https://doi.org/10.1145/564376.564421

Shetty, B. (2021, April 7). *An in-depth guide to how recommender systems work*. Built In. https://builtin.com/data-science/recommender-systems

Shu, J., Shen, X., Liu, H., Yi, B., & Zhang, Z. (2018). A content-based recommendation algorithm for learning resources. *Multimedia Tools and Applications*, *24*, 163–173. https://doi.org/10.1007/s00530-017-0539-8

Smyth, B., & Cotter, P. (2000). A personalized TV listings service for the digital TV age. *Knowledge-Based Systems*, *13*(2-3), 53–59. https://doi.org/10.1016/S0950-7051(00)00046-0

Suglia, A., Greco, C., Musto, C., de Gemmis, M., Lops, P., & Semeraro, G. (2017). A deep architecture for content-based recommendations exploiting recurrent neural networks. In M. Bielikova & E. Herder (Eds.), *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization* (pp. 202-212). Association for Computing Machinery. https://doi.org/10.1145/3079628.3079684

Sun, Y., Ji, Y., Zhu, H., Zhuang, F., He, Q., & Xiong, H. (2025). Market-aware long-term job skill recommendation with explainable deep reinforcement learning. *ACM Transactions on Information Systems (TOIS)*, *43*(2), Article 46. https://doi.org/10.1145/3704998

Wilson, D. C., Smyth, B., & O'Sullivan, D. (2003). Sparsity reduction in collaborative recommendation: A case-based approach. *International Journal of Pattern Recognition and Artificial Intelligence*, *17*(5), 863–884. https://doi.org/10.1142/S0218001403002678

Zarzour, H., Al-Sharif, Z., Al-Ayyoub, M., & Jararweh, Y. (2018). A new collaborative filtering recommendation algorithm based on dimensionality reduction and clustering techniques. In A. Shatnawi (Eds.), *9th International Conference on Information and Communication Systems* (pp. 102-106). IEEE. https://doi.org/10.1109/IACS.2018.8355449