A Novel Approach for Product Recommendation Using Smartphone Sensor Data

https://doi.org/10.3991/ijim.v16i16.31617

Moontaha Nishat Chowdhury(✉), H M Zabir Haque, Kazi Taqi Tahmid,
Fatema-Tuz-Zohora Salma, Nafisa Ahmed
Ahsanullah University of Science and Technology, Dhaka, Bangladesh
170104099@aust.edu

Abstract—Human Activity-based studies have become an omnipresent research topic in Machine Learning. Considering the countless impacts of human activity on persons' everyday life, we have analyzed the correlation between human activity and their product preferences in our study and proposed that daily human activity could be a metric for product recommendation models. To address this previously unaccounted phenomenon, a new approach is presented in our study that gives real-time recommendations to users by observing their activeness in daily life. However, product recommendation systems mostly believe in ratings, and the purchase behavior of users instead of investigating the precious insights of users' daily activities. But we examined smartphones' GPS sensor data using machine learning algorithms to urge insights from users' daily activeness and proposed a model for predicting the product of interest of the purchasers, based on the activeness of their daily life. Moreover, based on our model, we have introduced a prototype of a real-time recommendation system, especially for the retail shops that rely on users' implicit data from smartphone sensors to form product recommendations. For conducting our study, we developed an android application that—collects embedded smartphone sensor data and can detect objects to provide product recommendations and product details. Experiment shows, that our proffered daily activeness-based recommendation system using smartphone sensor data, performs with a precision of 66%, but it is also a promising performance because it does not use customers' explicit feedback.

Keywords—human activity, smartphone sensors, preferences, object detection, recommendations

1 Introduction

The proliferation of smartphones introduces golden opportunities in human activity pattern analysis. One of the reasons is that smartphones are usually equipped with sensors that can be used to infer the User's lifestyle pattern. Previously, constructing a dataset of sensor values was a complex task as it required wearable IoT devices for assembling the sensor values. But, the increasing use of smartphones and their embedded sensor system has enabled researchers to gather sensor data effortlessly and can capture any physical change in the environment. Smartphones' motion sensors like-accelerometer, gyroscope, magnetometer, and GPS sensors can record human mobility or motion [1], [2], [3], [4].

Prior studies have analyzed smartphone sensor data for profiling human behavior by applying machine learning techniques. Profiling human behavior means observing human behavior data and getting a pattern or structure from that data. For example—the group of individuals who exercise regularly is healthier than others. Here, these groups of healthy people may have a similar lifestyle pattern, and profiling can be done based on their daily exercise rate [5]. Former studies have developed Android smartphone applications to record longitudinal motion, location, and microphone sensor data and applied feature extraction techniques to raw sensor data to generate behavior profiles from monitoring patients' behavior [6] to predicting psychological developments [7]. Moreover, based on the sensor data of mobile phones, transport mode detection is done to know the User's life pattern or behavior [8], [9], [10]. An android application named "Carat" was developed to experiment with how it affects users' behavior in case of long-term use [11]. Surveys found out that among the beginner and advanced users of "Carat", the advanced users gradually learn to manage their battery and find the applications that consume more battery and replace them with alternatives.

Motivated by the impact of smartphone's embedded sensor data for human profiling or behavioral analysis, we have used smartphone sensor data to evaluate the correlation between daily activity and the product of interest of an individual. Additionally, we analyzed the customer's preferences for products based on their activeness in daily life. For that purpose, we have built android application that collects embedded smartphone sensor data in the background and gives product recommendation, shows rating on a product based on users' product scanning and rating history. From the latitude and longitude of the GPS sensor, we calculated the traveled distance of customers and applied unsupervised machine learning algorithms to cluster customers with similar patterns of traveled distance. Besides, we analyzed the product of interest among customers of each cluster. After mapping the outcome of interest and activeness, our model found a pattern where similar traveled distance customers have similar kinds of product interest. From this inference that there is a correlation between users' daily activity and product of interest, we have implemented a real-time Recommendation System (RS) based on daily activity.

Now the question is, why do we need RS? In today's era, retail shops or online shops or websites have really large catalogs of products. RS plays a vital role as there are users who know what they need specifically or looking for, whereas others face ambiguity while deciding what to pick from such a vast library of resources [12], [13], [14]. Recommendation engines or systems are the tools that are used to provide recommendations to users according to their product of interest [15], [16], [17]. Former work distinguishes RS's filtering techniques into four classes. These are demographic, Content-Based Filtering (CBF), collaborative filtering (CF), and hybrid procedures. Collaborative filtration is the most comprehensively used progress to scheme recommender systems [16], [18] and plays a vital role in the suggestion procedure [15], [17], [19], [18]. CBF algorithms attempt to suggest items to users based on the characteristics of the correspondence that the User formally chooses. Recommendations are founded on users' demographic profiles in demographic filtration. According to demographic parameters, the information given by the User is supposed to be the same like nationality or location, gender, age, etc., to recommend preferable products to the User. To beat some general problems occurring by the collaborative and content-based filtering procedures like cold start problem, overspecialization problem, and sparsity problem, the hybrid filtering procedure is inaugurated [19], [12] by combining multiple filtering techniques. Besides the filtering mentioned above methods, recently, the Next Basket Recommendation (NBR) techniques have been used widely in e-commerce and grocery shopping [20]. The goal of a next basket recommendation system is to recommend items for the next basket for a user [20], [21]. According to authors at [22], dividing users' shopping process into four categories, such as basket information, product sentiment, purchased items records, and click or viewed product records, gives promising results in the case of CBF. Similarly, as stated in [23] used the same process and considered three categories for information extraction from the user profile. An internet-based intelligent recommendation was proposed by authors at [24] where the hybrid approach was applied by combining content-based filtering and collaborative filtering, and a similarity measure of that system was done by using cosine similarity, Pearson correlation, and naive Bayesian classifier.

For each recommendation technique mentioned above, numerous researchers have enriched recommender algorithms that exert both explicit and implicit user feedback to enhance the outcome of a recommender system. Explicit feedback includes a user's commodity ratings and reviews. On the other hand, implicit feedback includes purchase histories, search histories, users' view or click patterns, etc. However, explicit feedback is not always accessible and is insufficient most of the time. Because most of the users are less likely to give feedback on products after using or purchasing. Again, much prior research has been done separately on human activity-based studies and personalized recommendation techniques. Research has not yet determined the correlation between users' daily activeness and product of interest. Moreover, most prior recommendation techniques use explicit feedback, such as ratings or reviews from users. Here, our proposed real-time recommendation system uses only implicit data from the User, where the implicit parameter is the activeness score. The prior studies did not investigate this metric for the recommendation system.

Using the implicit data of users' daily activeness, we have presented a new method for a real-time recommendation for retail shops using the hybrid recommendation technique. We aim to develop a framework to investigate the correlation between users' daily traveled distance and product of interest. The Implicit data is collected from users' embedded smartphone sensor GPS data. From this sensor data, users are be segmented into different groups as discussed above. Each group of people holds an activeness tag and score based on their daily activeness, such as Active, Moderately Active, and Less Active. According to the activeness tags, these activity scores are used as the implicit data in our proposed recommender system.

2 Proposed architecture

A detailed description of our approach is discussed in this section. Firstly, an overview of the study is given, followed by details of the step-by-step implementation of the real-time recommendation system by analyzing the correlation between users' daily activeness and product of interest. The main steps of this section are illustrated in Figure 1.

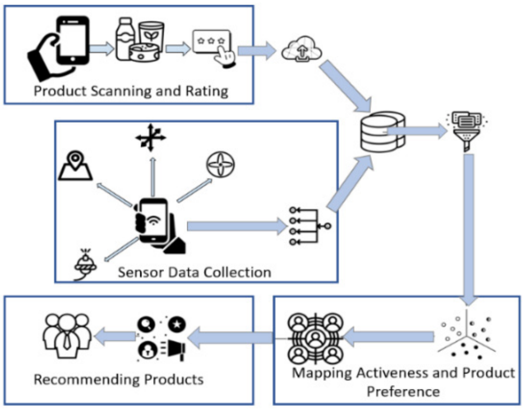

Fig. 1. Proposed system architecture

We have built an Android Application for collecting users' preferences and embedded smartphone sensor data. The application can store different sensor values like Accelerometer, Gyroscope, Magnetometer, GPS, light, pressure sensor values, etc. To observe users' preferences, we have introduced our app's product scanning and rating facilities. In order to make it capable of scanning products, we have chosen some specific products for our study and integrated object detection facilities in our android application. These products are selected by us carefully in order to distinguish between users' interests. We have taken products from 3 categories and in each category, there are 2 to 3 variations in product type. The list of products we have used is exhibited below in Table 1. We have trained our model using TensorFlow Lite with product images captured by us. For every product, we have captured images from different angles and made a dataset of images containing 40 to 50 images per product in order to train the model. So as to achieve our model for object detection, we have used the Teachable Machine Learning tool. For analyzing a user's activity pattern, we have estimated the total distance traveled by the User during a specific period by using the GPS sensor values.

Table 1. List of products used for object detection and classification

Product Type	**Soap**	**Hand Sanitizer**	**Washing Detergents**	**Organic Food**	**Processed Food**	**Wrist Watch**	**Sunglass**
Product Category	Category1 Variant 1 (Cat1V1)	Category1 Variant 2 (Cat1V2)	Category1 Variant 3 (Cat1V3)	Category2 Variant 1 (Cat2V1)	Category2 Variant 2 (Cat2V2)	Category3 Variant 1 (Cat3V1)	Category3 Variant 2 (Cat3V2)

2.1 Data acquisition

Sensor data collection. As our goal is to acknowledge the behavior of users in their daily life, we have collected sensor data using our smartphone's Android application. For that purpose, we have collected data of 53 volunteers. For smartphone sensor data, volunteers just have to keep our application running in the background. This application stores the sensor data including Accelerometer, Gyroscope, Magnetometer, GPS, light sensor, etc. temporarily in local storage of smartphones using SQLite database system. After a certain amount of time, data stored in SQLite gets transferred to the Firebase Realtime database under the User's unique user Id. These sensor data are then collected from firebase and converted into CSV files for further computations.

Users' interest collection. In addition to sensor data, we have also collected users' interests in different products. In order to gather Users' interest, we have introduced **Product Scanning and Rating Facilities** in our app. A user can scan a specific product, see details and the current rating of that product, and also give a rating for that particular product. With the aim of aggregating users' interest, we have assumed that the User has an interest in products that are being scanned or rated by him/her. Therefore, each time a user scans or gives a rating to a particular product, we are expecting that the User has an interest in that product and an entry in the Firebase Realtime database has been stored by the user Id and Product Id of that product.

2.2 Users' movement estimation

For the purpose of estimating users' movement, we have constructed our dataset stored in Firebase collecting sensor values (Accelerometer, Magnetometer, Gyroscope, GPS, Light) from our Android App. As the data is in raw form and if we had sent these data to our model, this would have caused miscalculations. Hence, it would have directly influenced the capability of our model to learn. For that reason, before feeding data into our model, we have transformed the raw data into cleaned data for the analysis using data preprocessing techniques. For our experiment, we have used only GPS Sensor data by extracting other sensor values as those values expand the complexity. We took into consideration GPS sensor data of 14 days and have taken columns of User Id, timestamp, longitude, and latitude. As a part of preprocessing, we have also extracted some rows of duplicate longitude and latitude values from the dataset. We have kept only distinctive data for the flexibility of the analysis in our experiment. After that, the Haversine formula has been used for the User's traveled distance calculation. We have calculated the distance between each consecutive point and summed up all the distance values for a particular user Id and also found the maximum value of each User.

Clustered users traveled distance. A k-means clustering algorithm was performed to cluster the users based on their traveled distance. The Elbow method was used to find the number of centroids (k).

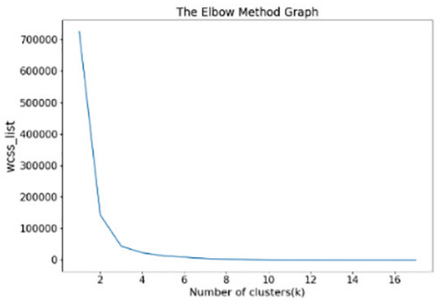

Fig. 2. Determining the number of k using the elbow method

Having implemented the elbow method (Figure 2), k=3 was taken for implementing the clustering method. Hence, 53 customers are grouped into 3 clusters based on users' traveled distance after the clustering procedure.

2.3 Customer segmentation

Having collected data on users' interest in different products from our smartphone's Android application, we have prepared the dataset of customer preferences for the customer segmentation process. Here in the dataset, Cat1V1 represents variant 1 of category 1 which represents the soap type of product from the Table 1 in Figure 3. We have prepared a dataset of 53 customers.

User ID	Cat1V1	Cat1V2	Cat1V3	Cat2V1	Cat2V2	Cat3V1	Cat3V2
1	1	1	0	0	1	0	0
2	1	0	1	1	0	0	1
3	1	0	0	1	1	0	1
4	1	1	0	0	0	0	0

Fig. 3. Snapshot customer preference dataset

After that, a k-means clustering algorithm has been performed for clustering customers based on their preferences in product choices. The elbow method has been used for deciding the number of centroids (k). After performing the elbow method, k=3 was taken for performing the clustering method. Hence, customers are grouped into 3 clusters after the clustering process.

2.4 Real-time recommendation engine construction

The mechanism of how we will segment customers based on their daily activity using embedded smartphone sensor data has been discussed above. Along with collecting smartphone sensor data and ratings, and providing product details, our developed android application also provides real-time recommendations to the users. This section illustrates the methodology of how we have developed a Recommendation Engine based on users' activity.

Implicit feedback for personalized recommendation. For serving real-time personalized recommendations to the User, our study has focused on customers' implicit data. Primarily we have collected implicit data from users' embedded smartphone sensor data. After doing customer segmentation based on daily activeness, users of three different characteristics have been detected: Active, Moderately Active, and Less Active. Users with a high distance traveled within 14 days [25] are clustered as Active. On the other hand, users with very low distance traveling records are Less Active, and those who are in between these two extreme categories are clustered as Moderately Active. According to these three tags we have given each category user an activity score. Users with Active tags have activity scores of 50, Moderate actives are scored 30 and less active tag users have activity scores of 10. These activity scores according to the activeness tags are used as the implicit data in our proposed recommender system. To walk through the features of our product recommendation app using sensor data, we use the following toy example.

X is a smartphone user who has participated as a volunteer in our research. He obtained an activeness score of 50 and is currently scanning "Hand Sanitizer A" (Figure 4). Now based on X's Scan product record and activeness score record our application recommends a product within the same category which is "Hand Sanitizer B" and is heavily scanned by other customers with an activeness score of 50.

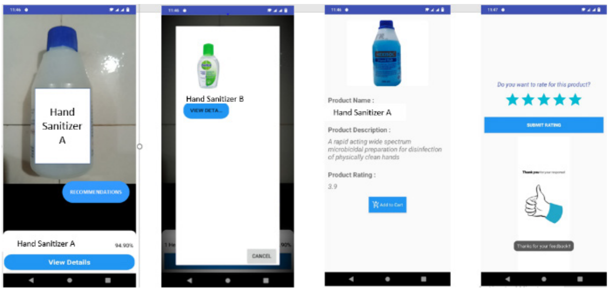

Fig. 4. Snapshot of our android application's real-time recommendation (Hand sanitizer)

Personalized hybrid recommender. Over the years, a large number of RS have utilized the hybrid techniques to overcome the shortcomings of only using collaborative filtering or content-based filtering to serve recommendations according to users' preferences. With a view to conducting our research on predicting products of interest by observing users' preferences and smartphone sensor data, we have focused on users' activity scores and scanned products record for CBF and CF approaches respectively. To address these previously unaccounted phenomena, similar people having the same daily activity have also similar product choices, a new approach is presented in which we have used the activity score as implicit data to measure similarity among users, and for content-based filtering, we have built user profile based on scanned items by users.

Both these methods face problems when a new user arrives. To overcome this problem, a common approach is a popularity-based recommendation, where the highly-rated products or heavily rated products are suggested to the new User. Contrary to similar methods which make use of ratings, our method does not require a highly rated or heavily rated product list, rather we have used highly scanned products record to overcome the cold-start problem of new users. In this study, we assume that products which are heavily scanned are more popular than other products among the product category, and recommendations for a new user will be given with the highly scanned product according to the category.

Algorithm				
Input	:	U	←	set of users with corresponding activeness score, scanned product list, purchased product list
		P	←	set of products
Output	:	R	←	Output real-time recommendation
Step1	:	R	←	Ø
Step2	:			if activity < 14 days then
Step3	:			for $u \in$ U do
Step4	:	PRP	←	Predicted Product interest based on heavily scanned or popularity-based approach
Step5	:	R	←	R ∪ PRP
Step6	:			end for
Step7	:			else
Step8	:	Sim (u,u')	←	cosine similarity score between u and u' according to their activeness score
Step9	:			if sim ((u, u) > (threshold = 0.65)
Step10	:	PRC	←	Predicted Product interest based on collaborative filtering
Step11	:	UP	←	Generated individual user profiles based on scanned product data
		PRCB	←	Predicted Product interest based on content-based filtering using user profile UP
Step12	:	R	←	R ∩ PRC ∩ PRCB

User-based collaborative filtering. In traditional collaborative filtering-based approaches, explicit feedback, such as ratings or reviews is used to measure the similarity between users or items. Here we have done user-based collative filtering using implicit data of users which is activity score. This has been done from the assumption that people with similarly traveled activities, also have similar products of interest. Hence, by recording the purchased items history of users, we have built the utility matrix of user-item (Figure 5) with corresponding users' activeness scores. For example, if a user bought 3 items A, D, G (among A, B, C, D, E, F, G) and that User is moderately active in his or her daily life. According to our previously mentioned scheme, that User holds an activity score of 30. Now, in the utility matrix, the User will contain 30 against the purchased items A, D, and G, and for other items, he did not buy yet will obtain a zero value.

ProductName User ID	Cat1V1	Cat1V2	Cat1V3	Cat2V1	Cat2V2	Cat3V1	Cat3V2
43	10.0	0.0	0.0	0.0	10.0	0.0	10.0
44	30.0	0.0	0.0	0.0	30.0	0.0	0.0
45	0.0	30.0	0.0	0.0	30.0	0.0	30.0
46	10.0	0.0	0.0	0.0	10.0	0.0	10.0
47	50.0	0.0	0.0	50.0	0.0	0.0	50.0

Fig. 5. A snapshot of utility matrix with user activeness score

This utility matrix is then used to calculate the cosine similarity between two users. After the similarity measure, we have done thresholding, rather than just recommending products based on "Top K" similar to users like the traditional approach. For a user, u when the similarity score with other users (*u*) is greater than or equals T=0.65, then those users are considered similar to the User u. Finally, according to a similar user, the products that have not yet been explored by User u have been predicted for the recommendation. As in our proposed approach, we use users' activity score which is fixed for a particular user, such as 50 for the active users, 30 for moderately active users, and 10 for the less active users. Hence, this system is not considering an item that is bought by a user and the User did not like that item. This case is handled by the traditional or other prior collaborative filtering approaches as they use the explicit feedback ratings on items and negative ratings or very low values in a rating scale (i.e., rating with 1 start in a range of 1 to 5 stars) are also considered while calculating similarity score. Hence, to overcome this shortcoming, we have also done content-based filtering on users' profiles.

Content-based filtering. For performing content-based filtering, in order to build a user profile for each User, we have considered the products which are scanned by that particular User. From the scanned product list of a user, we take the category id and variant id, in order to get the item attribute. From this item attribute, the content type which might be liked by the User can be achieved and a user profile is built. Furthermore, we also take into account the scan count of a specific attribute (category id + variant id) by a user.

To determine item attributes to the scanned item list, category id and variant id of users' scanned product are marked as 1, otherwise 0. After that, the user profile is generated by taking the dot product of the transpose of the User's scanned product attribute and counting the scanned products by the User. After generating the user profile for each User, the predicted items for recommendation are calculated by taking the dot products of all the items attributed to users' profiles, and a weighted average is taken. After building the collaborative filtering-based recommender model and content-based filtering recommender model separately, we have merged these two recommendations to get more accurate recommendations for our android application users, who have been using our application for more than 2 weeks. Users who have been using our application for less than 2 weeks or have not scanned any product will receive a popularity-based recommendation, where we are measuring popularity against high scanned products.

3 Results and discussions

The purpose of the study was to analyze the correlation between the daily activeness of users and the product of preferences and provide real-time personalized recommendations. This section comprises two sub-sections. The first sub-section depicts the experimental result of clustering users based on embedded smartphone sensor data and product of interest and, the second sub-section discusses the performance of our proposed recommendation system.

3.1 Analysis of traveled distance and product of interest

In this study, we calculated the traveled distance using the haversine formula from the GPS longitude and latitude data of users. Based on the traveled distance, we got three categories of users. In Figure 6, the x-axis denotes the distance traveled in meters and the y-axis denotes speed. The 3 clusters formed based on the 14 days' travel record of volunteers have traveled mostly below 10,000 m, between 10,000 to 30,000 m, and above 40,000 m.

Fig. 6. Cluster results of customers based on traveled distance

There are very few customers who have traveled more than 60,000 m on average and most of the volunteers' average travel distance is under 30,000 m (Figure 7). By analyzing the value of each cluster, we labeled these three groups into Active, Less Active, and Moderate Active.

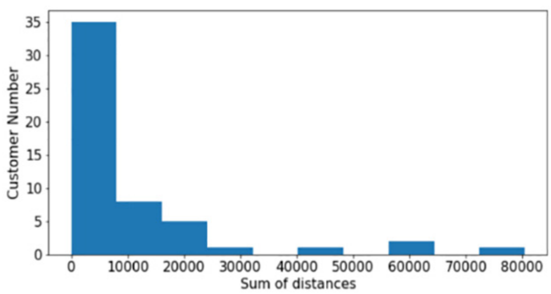

Fig. 7. Frequency distribution of the sum of the distance of customers

After clustering users based on traveled distance, we have analyzed the product of interest of users of each cluster tagged as Active, Moderate Active, and Less Active. From Figure 8, it can be seen that there exists a product choice similarity among users tagged as 'Moderate active' or 'Active'. Users tagged as 'Moderately Active' are mostly interested in 'Cat1V2', 'Cat2V2', and 'Cat3V2' products. On the other hand, the highest scanned product of 'Active' users are 'Cat1V2', and 'Cat2V2'. Moreover, another fact is noticed from the following heatmap which is, that the 'Cat3V1' product is not at all scanned by the Moderate Active users, whereas, very few active people are interested in that product.

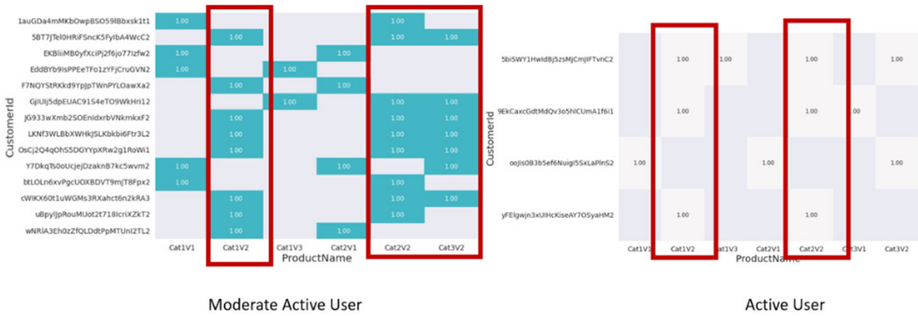

Fig. 8. Heatmap of product preferences of customers tagged as moderate active and active

Hence, Figure 8 shows there exists a pattern in the product of interest among people having activeness-based similarity.

3.2 Performance of the proposed real-time recommendation system

Our Proposed user activeness-based similarity score measure gives a promising performance. As shown in Figure 9, the darker the color the higher the similarity among users, and also hold similar tags activeness (Active, Moderate active, less active) in most of the cases (Figure 4.1).

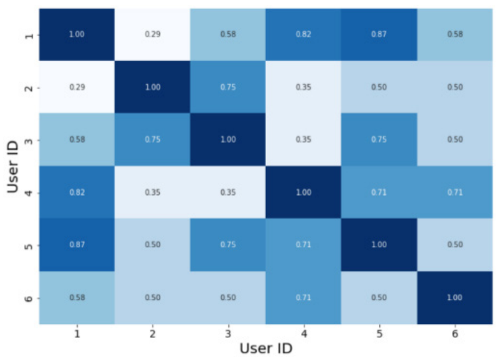

Fig. 9. Heatmap of similarity scores between users based on activeness score

To evaluate the result of our real-time recommendation system, we have compared the prediction of our proposed collaborative filtering model with traditional collaborative filtering' prediction. To evaluate our recommendation system, we focus on the accuracy of the true product of interest or suggestion by the model and emphasize the precision of the RS. Here, our main concern is how accurate our model is in predicting the true product choices of customers. In the case of traditional rating-based collaborative filtering results in higher precision, whereas inactivity score-based collaborative filtering scores in 66%. This is because in the activeness-based model we are only using a specific activeness score for the corresponding User based on his or her purchased history. On the other hand, it also considers the negative rating to compute the similarity between users in a rating-based model. The below Table 2 shows a sample of recommended product lists of user id 1 & 5 using rating-based collaborative filtering and activeness score based collaborative filtering. Number of products suggested varies based on users scanned and rated product data.

Table 2. Sample comparison of the prediction for the traditional collaborative filtering of user ID 1 & 5 with our proposed model

User ID	Rating Based Collaborative Filtering	Activeness Score Based Collaborative Filtering
1	Soap (Lifebuoy)	Soap (Lifebuoy)
	Hand Sanitizer (Hexisol)	Breakfast item (Saad Atta) [negative rating was given to this product]
5	Sunglasses	
	Hand Sanitizer (Hexisol)	Hand Sanitizer (Hexisol)
	Breakfast item (Processed Food)	Breakfast item (Processed Food)

In the case of the product of interest prediction, using users' activeness score-based matrix also gave almost similar but less accurate results in the case of only using collaborative filtering. To overcome this shortcoming of not considering the negative rating we have combined the scan-based content-based filtering and get a more reliable real-time recommendation. Together these results provide important insights into that, people with similar kinds of activity also have similar products of interest.

4 Conclusion

The results of this research support the idea that human activity can be a parameter in the case of customer segmentation and thus segmenting customers based on activeness can play a vital role in predicting customer behavior and preference. This study discovers similar patterns in product selection based on their measurement of activeness and gives real-time recommendations to the users accordingly. This study has shown that by observing users' GPS value we can cluster them into different groups and label them into different groups on the parameter of activeness. We have captured the embedded sensor data of smartphones of our volunteers for 2 weeks and estimated their traveled distance to identify their activeness in daily life. The research has also shown that an active person's choices and preferences differ from a less active person's choices and preferences. Thus, marketing policies and strategies to reach out to targeted customers can be done smartly and comfortably. Based on this hypothesis, our proposed recommendation system—real-time recommendations with user activeness; considering this as an implicit parameter also gives promising results.

The primary goal of our study was to observe users' smartphone sensor data and products of interest and investigate the correlation between these two. The results show that there exist similar product choices within the same clusters of people. Because of our impotence in collecting data for the ongoing Covid-19 pandemic, we cannot deploy our application and collect data from a wide range of users. Thus, we have conducted our study only on 53 volunteers and collected their embedded sensor data. This small number of samples lead to some shortcomings. Firstly, small number of volunteers creates biasness in number of active, moderate active and less active people. From our 53 volunteers GPS sensor data only 25% are clustered as active person. Due to the huge imbalance in active, moderate active and less active peoples' number, our model could not get the actual pattern in product choice. Secondly, we could not gather the actual customer preferences dataset for the same reason. The rated or scanned product record is limited in 53 people. Hence, the model could not get much variation in their product choice but which might be occur in case of large scale of data. Owing to this problem, our model could not capture the actual preferences of the customers.

Our research studies propose a pipeline for correlating human activeness with his/her personal preferences, and thus, we believe that with the actual dataset, our model will be able to capture a more precise correlation between human activeness and preferences. However, in the future, we aim to work with the real dataset of sensor values and customer preferences to generate a more stable model for our research. Furthermore, with our current proposed recommendation model, we aim to integrate the demographic filtering technique based on location data to recommend the topmost trending products of that particular location to the users.

5 Acknowledgement

Researchers greatly appreciate the 53 volunteers who have contributed to data collection and analysis in this research.

6 References

[1] Bayat, A., Pomplun, M., & Tran, D. A., "A Study on Human Activity Recognition Using Accelerometer Data from Smartphones," *Procedia Computer Science,* vol. 34, pp. 450–457, 2014. https://doi.org/10.1016/j.procs.2014.07.009

[2] Kwapisz, J. R., Weiss, G. M., & Moore, S. A., "Activity Recognition Using Cell Phone Accelerometers," *ACM SigKDD Explorations Newsletter,* vol. 12, pp. 74–82, 2011. https://doi.org/10.1145/1964897.1964918

[3] Hemminki, Samuli, Nurmi, Petteri, Tarkoma, & Sasu., "Accelerometer-Based Transportation Mode Detection on Smartphones," 2013. https://doi.org/10.1145/2517351.2517367

[4] Majumder, S., & Deen, M. J., "Smartphone Sensors for Health Monitoring and Diagnosis," *Sensors,* vol. 19, p. 2164, 2019. https://doi.org/10.3390/s19092164

[5] Kelly, D., Condell, J., Curran, K., & Caulfield, B., "A Multimodal Smartphone Sensor System for Behaviour Measurement and Health Status Inference," *Information Fusion,* vol. 53, pp. 43–54, 2020. https://doi.org/10.1016/j.inffus.2019.06.008
[6] Cheng, W.-Y. et al., "Human Activity Recognition from Sensor-Based Large-Scale Continuous Monitoring of Parkinson's Disease Patients," 2017. https://doi.org/10.1109/CHASE.2017.87
[7] Ahmed, Umair, Ali, Muhammad Faizyab, Javed, Kashif, Babri, & Haroon Atique, "Predicting Physiological Developments from Human Gait using Smartphone Sensor Data," *arXiv preprint arXiv:1712.07958,* 2017.
[8] Jahangiri, A., & Rakha, H. A., "Applying Machine Learning Techniques to Transportation Mode Recognition Using Mobile Phone Sensor Data," *IEEE Transactions on Intelligent Transportation Systems,* vol. 16, pp. 2406–2417, 2015. https://doi.org/10.1109/TITS.2015.2405759
[9] Wang, S., Chen, C., & Ma, J., "Accelerometer Based Transportation Mode Recognition on Mobile Phones," 2010. https://doi.org/10.1109/APWCS.2010.18
[10] Shafique, M. A., & Hato, E., "Use of Acceleration Data for Transportation Mode Prediction," *Transportation,* vol. 42, pp. 163–188, 2015. https://doi.org/10.1007/s11116-014-9541-6
[11] Athukorala, Kumaripaba, Lagerspetz, Eemil, Von Kügelgen, Maria, Jylhä, Antti, Oliner, Adam J, Tarkoma, Sasu, Jacucci, & Giulio, "How Carat Affects User Behavior: Implications for Mobile Battery Awareness Applications," 2014. https://doi.org/10.1145/2556288.2557271
[12] Dhelim, S., Ning, H., Aung, N., Huang, R., & Ma, J., "Personality-Aware Product Recommendation System Based on User Interests Mining and Metapath Discovery," *IEEE Transactions on Computational Social Systems,* vol. 8, pp. 86–98, 2020. https://doi.org/10.1109/TCSS.2020.3037040
[13] Geetha, G., Safa, M., Fancy, C., & Saranya, D., "A Hybrid Approach using Collaborative Filtering and Content Based Filtering for Recommender System," 2018. https://doi.org/10.1088/1742-6596/1000/1/012101
[14] Chen, L., Li, R., Liu, Y., Zhang, R., & Woodbridge, D. M. K., "Machine Learning-Based Product Recommendation Using Apache Spark," 2017. https://doi.org/10.1109/UIC-ATC.2017.8397470
[15] Ekstrand, M. D., Riedl, J. T., & Konstan, J. A., Collaborative Filtering Recommender Systems, Now Publishers Inc, 2011. https://doi.org/10.1561/9781601984432
[16] Konstan, J. A., & Riedl, J., "Recommender Systems: from Algorithms to User Experience," *User Modeling and User-Adapted Interaction,* vol. 22, pp. 101–123, 2012. https://doi.org/10.1007/s11257-011-9112-x
[17] Shi, Y., Larson, M., & Hanjalic, A., "Collaborative Filtering Beyond the User-Item Matrix: A Survey of the State of the Art and Future Challenges," *ACM Computing Surveys (CSUR),* vol. 47, pp. 1–45, 2014. https://doi.org/10.1145/2556270
[18] Ahn, H. J., "A New Similarity Measure for Collaborative Filtering to Alleviate the New User Cold-Starting Problem," *Information Sciences,* vol. 178, pp. 37–51, 2008. https://doi.org/10.1016/j.ins.2007.07.024
[19] Qian, G., Sural, S., Gu, Y., & Pramanik, S., "Similarity Between Euclidean and Cosine Angle Distance for Nearest Neighbor Queries," 2004. https://doi.org/10.1145/967900.968151
[20] Li, M., Jullien, S., Ariannezhad, M., & de Rijke, "A Next Basket Recommendation Reality Check," *arXiv preprint arXiv:2109.14233,* 2021.
[21] Hu, H., He, X., Gao, J., & Zhang, Z. L., "Modeling Personalized Item Frequency Information for Next-Basket Recommendation," 2020. https://doi.org/10.1145/3397271.3401066
[22] Lee, C. H., Kim, Y. H., & Rhee, P. K., "Web Personalization Expert with Combining Collaborative Filtering and Association Rule Mining Technique," *Expert Systems with Applications,* vol. 21, pp. 131–137, 2001. https://doi.org/10.1016/S0957-4174(01)00034-3

[23] Cho, Y. H., & Kim, J. K., "Application of Web Usage Mining and Product Taxonomy to Collaborative Recommendations in E-commerce," vol. 26, pp. 233–246, 2004. https://doi.org/10.1016/S0957-4174(03)00138-6
[24] Montaner, M., López, B., & De La Rosa, J. L., "A Taxonomy of Recommender Agents on the Internet," *Artificial Intelligence Review,* vol. 19, pp. 285–330, 2003. https://doi.org/10.1023/A:1022850703159
[25] Stanley, Kevin, Yoo, Eun-Hye, Paul, Tuhin, Bell, & Scott, "How Many Days Are Enough?: Capturing Routine Human Mobility," *International Journal of Geographical Information Science,* vol. 32, pp. 1485–1504, 2018. https://doi.org/10.1080/13658816.2018.1434888

7 Authors

Moontaha Nishat Chowdhury received her Bachelor's degree in Computer Science and Engineering from Ahsanullah University of Science and Technology, Dhaka, Bangladesh (email: 170104099@aust.edu).

H M Zabir Haque is an Assistant Professor at Ahsanullah University of Science and Technology, Dhaka, Bangladesh. He has received his Master of Science in Bioinformatics from the University of Saskatchewan, Canada, and a Bachelor's degree in Computer Science and Engineering from Ahsanullah University of Science and Technology, Dhaka, Bangladesh. His research interests include Bioinformatics, Computational Biology, and Machine Learning (email: zabir.haque.cse@aust.edu).

Kazi Taqi Tahmid received his Bachelor's degree in Computer Science and Engineering from Ahsanullah University of Science and Technology, Dhaka, Bangladesh (email: taqitahmid97@gmail.com).

Fatema-Tuz-Zohora Salma received her Bachelor's degree in Computer Science and Engineering from Ahsanullah University of Science and Technology, Dhaka, Bangladesh (email: zohorasalma@gmail.com).

Nafisa Ahmed received her Bachelor's degree in Computer Science and Engineering from Ahsanullah University of Science and Technology, Dhaka, Bangladesh (email: ahmednafisa11@gmail.com).

Article submitted 2022-04-11. Resubmitted 2022-07-13. Final acceptance 2022-07-25. Final version published as submitted by the authors.