Analyzing Mobile Device Usage and User Behavior Using KMeans Clustering

Jayasurya Murali

October 20, 2024

Abstract

This study employs the Knowledge Discovery in Databases (KDD) process to analyze mobile device usage and user behavior. Using KMeans clustering on various features, such as app usage time and battery drain, we uncover distinct user behavior patterns. This analysis provides insights that could help optimize mobile device performance and improve user experience.

1 Introduction

Mobile devices have become indispensable tools in modern society, influencing how individuals interact with technology. Understanding user behavior can lead to the development of better user-centric applications and device optimizations. This study uses the KDD methodology to analyze a comprehensive dataset of mobile device usage and user behavior. The primary goal is to identify patterns in user behavior through clustering techniques.

2 Methodology

2.1 KDD Process Overview

The Knowledge Discovery in Databases (KDD) process involves five key steps: Selection, Preprocessing, Transformation, Data Mining, and Interpretation. This study follows each phase to extract knowledge from the dataset and apply clustering techniques to uncover behavioral patterns.

2.2 Selection

The dataset used in this study is publicly available on Kaggle, titled *Mobile Device Usage and User Behavior Dataset*. The selected features for analysis include:

- App Usage Time (min/day)
- Screen On Time (hours/day)

- Battery Drain (mAh/day)
- Data Usage (MB/day)
- User Behavior Class

2.3 Preprocessing

In this phase, missing data were handled, and the numerical features were normalized to ensure consistent ranges across variables. The data was cleaned and prepared for transformation and clustering.

2.4 Transformation

Data transformation involved scaling the numerical variables using StandardScaler. This step ensured that all features were on a similar scale, which is important for distance-based clustering algorithms like KMeans.

2.5 Data Mining

KMeans clustering was employed to group users based on their mobile device usage patterns. Four clusters were identified, representing distinct behavioral patterns based on app usage time, screen on time, and battery drain.

App Usage (min/day)	**Screen On Time (hrs/day)**	**Battery Drain (mAh/day)**	**Data Usage (MB/day)**	**User Behavior Class**
393	6.4	1872	1122	4
268	4.7	1331	944	3
154	4.0	761	322	2
239	4.8	1676	871	3
187	4.3	1367	988	3

Table 1: Sample Data from Mobile Device Usage and User Behavior Dataset

3 Results and Discussion

The clustering results revealed four distinct user behavior patterns. Cluster 1 represented users with high app usage and high screen on time, while Cluster 4 represented users with minimal app usage and low battery drain. The clusters align with the original user behavior class labels, demonstrating the efficacy of the KMeans algorithm in identifying meaningful behavioral segments.

The analysis highlights how user behavior varies based on device usage, suggesting opportunities for optimizing mobile devices for different user groups. For instance, high-usage

clusters may benefit from battery optimization features, while low-usage clusters may require different forms of user engagement.

4 Conclusion

This study successfully applied the KDD process to analyze mobile device usage and user behavior data. By using KMeans clustering, distinct user behavior patterns were identified, providing actionable insights for mobile device manufacturers and app developers. Future work can explore more advanced clustering techniques or apply this analysis to other domains of user behavior analytics.