ISSN 1816-6075 (Print), 1818-0523 (Online)
Journal of System and Management Sciences
Vol. 15 (2025) No. 2, pp. 290-308
DOI:10.33168/JSMS.2025.0218

Consumer Behavior Analysis in Gamified Mobile Banking: Clustering and Classifier Evaluation

Putri Taqwa Prasetyaningrum[1,2,] Purwanto Purwanto[1,3], Adian Fatchur Rochim[4]

[1] Doctoral Program of Information Systems, Universitas Diponegoro, Semarang, Indonesia

[2] Faculty of Information Technology, Universitas Mercu Buana Yogyakarta,, Indonesia

[3] Department of Chemical Engineering, Faculty of Engineering, Diponegoro University, Semarang

[4]Faculty of Engineering, Universitas Diponegoro, Semarang, Indonesia

putritaqwa@students.undip.ac.id

Abstract. This study explores the application of K-Means clustering and various machine learning classifiers to analyze consumer behavior in gamified mobile banking applications. Using a survey of 451 active mobile banking users in Indonesia, the research investigates user interactions with gamification elements and preferences within the applications. The K-Means algorithm is employed to segment users into distinct clusters based on their engagement patterns, transaction frequencies, and usage behaviors, offering insights into diverse consumer groups. The optimal number of clusters is determined through the elbow method and silhouette score analysis. Subsequently, the performance of several classifiers, including Support Vector Machine (SVM), Random Forest, AdaBoost, Decision Tree, Logistic Regression, and Gradient Boosting, is evaluated in predicting cluster assignments. Comprehensive metrics such as accuracy, precision, sensitivity, specificity, F1 score, Cohen's kappa, and Matthews correlation coefficient are utilized for rigorous assessment. The results reveal the SVM classifier's superior and consistent performance across various metrics, indicating its suitability for modeling consumer behavior in gamified mobile banking environments. By uncovering distinct user segments and predicting their behaviors, this study provides a data-driven foundation for personalizing user experiences, enhancing gamification strategies, and tailoring marketing efforts within the mobile banking sector.

Keywords: K-Means Clustering, Consumer Behavior, Mobile Banking Applications, Gamification, Machine Learning Classifiers, Performance Evaluation, SVM Classifier, Cluster Validation

1. Introduction

In the banking industry, digitization has become a top priority to enhance convenience, efficiency, and customer satisfaction while reducing operational costs. Gamification elements in online banking are increasingly used to influence customer behavior and engagement, transitioning from conventional to digital platforms (Rodrigues et al., 2016). Previous studies have shown that gamification strategies can significantly enhance user interaction and participation in non-game contexts (Martín-Gutiérrez et al., 2017).

Gamification in mobile banking applications has been shown to enhance user engagement and interaction. For instance, (Huang & Soman, 2013) found that gamification elements in web design can significantly influence expected customer behavior. (Martín-Gutiérrez et al., 2017) demonstrated how gamification in marketing strategies increases user participation and interaction. These studies provide a foundation for understanding the potential benefits of gamification in mobile banking. With the growing number of gamification system implementations (Koivisto & Hamari, 2019). and significant investments are expected in this area (Technavio, 2019). However, there are still shortcomings in understanding the best ways to motivate and engage users in gamified systems (Koivisto & Hamari, 2019). In previous research, games have long been recognized as an engaging form of interaction that can lead to addiction (Granic et al., 2014; Grüsser et al., 2007; Hamari, 2017; Mallon & Lynch, 2014). Therefore, it is not surprising that marketing practitioners are beginning to look at games' potential to enhance customer engagement (Hofacker et al., 2016; Huotari & Hamari, 2017). (Wolf et al., 2020) Finding that services leveraging gamification help companies prioritize experiences crucial to their business goals. The trend of transforming services and products to resemble games has become one of the main directions in recent technology (Huotari & Hamari, 2017; Koivisto & Hamari, 2019). Gamification has also been utilized for employee education across various industries, fostering brand engagement and driving positive behavior change (Nancy V. Wünderlich Paderborn et al., 2019). This study uses a K-means cluster analysis to distinguish different user groups based on their behavior within a gamified fitness program (Krath et al., 2022).

Understanding consumer behavior in gamified environments is crucial, particularly in mobile banking applications where personalized experiences can drive user engagement and satisfaction. This study aims to analyze consumer behavior within gamified mobile banking applications, focusing on how different gamification elements influence user interactions and preferences.

Personalization in mobile banking applications is essential for enhancing user engagement and satisfaction. By integrating gamification strategies with personalized user experiences, banks can create more engaging and beneficial interactions for their customers. This study leverages K-Means clustering and machine learning classifiers to identify distinct user segments, providing insights for developing targeted gamification strategies (M. Ahmed et al., 2020).

Despite the growing interest in gamification, there is a lack of focused research on its impact within mobile banking applications. Existing literature broadly covers gamification's benefits but often overlooks specific consumer behavior analysis in this context. This study addresses these gaps by combining K-Means clustering and machine learning classifiers to analyze user behavior in a gamified mobile banking environment, offering new insights into user engagement and satisfaction.

The rapid adoption of mobile banking and the increasing incorporation of gamification strategies underscore the urgency of this research. Financial institutions are striving to create more personalized and engaging user experiences. Understanding the behavioral patterns and preferences of customers is crucial for designing effective gamification strategies that enhance user satisfaction and loyalty.

In Srivastava's research, hierarchical segmentation techniques proved effective in categorizing consumers by considering factors such as age and income. Further segmentation was applied to convenience, entertainment, and engagement to tailor gamification marketing strategies. This indicates that demographic aspects and consumer preferences play a crucial role in the success of gamification, providing marketers with new insights for more targeted marketing strategies (Srivastava & Bag, 2021).

Modern marketing requires careful integration between customer segmentation and churn prediction modeling. Although this process is complex—due to involving analysis of heterogeneous customer behavior data—recent research suggests that a hybrid approach in segmentation can enhance churn prediction accuracy. This innovation combines suitable distance measurement techniques for numerical and categorical data, facilitating more precise segmentation before churn classification, which has been shown to improve model performance in the mobile gaming industry (Perišić & Pahor, 2023). The selection of K-Means clustering is due to its simplicity, scalability, and effectiveness in handling large datasets, which are highly advantageous for analyzing consumer behavior in mobile banking. Its centroid-based approach excels in grouping users based on their interaction patterns with gamified applications, offering clear and actionable insights. The suitability of this method with the data characteristics—dimensions and distributions of user behaviors—in this study is confirmed by initial analyses and supported by existing literature, thus providing a strong foundation for its selection over other clustering techniques.

This approach aims to influence user behavior, encourage regular use, and improve overall satisfaction. Despite the growing interest in gamification, there is a need for focused research on its impact specifically within mobile banking applications. While existing literature broadly covers gamification's benefits in various sectors, studies directly addressing consumer behavior analysis in gamified mobile banking environments remain limited. This gap highlights the necessity of examining how gamification influences user interactions and preferences in this context. This study aims to address the following research questions: How does gamification affect consumer behavior in mobile banking applications? What are the distinct user behavior clusters in a gamified mobile banking environment? How effective are various machine learning classifiers in predicting user cluster assignments based on their interactions with gamification features? By answering these questions, this research seeks to provide valuable insights into user engagement strategies for financial institutions, contributing to both academic discussions and practical applications in digital banking. While gamification has been extensively studied in general terms, specific gaps in the literature include a lack of focused studies on consumer behavior analysis in gamified mobile banking applications, limited exploration of the combined use of clustering algorithms and machine learning classifiers to understand user behavior in this context, and insufficient examination of the effectiveness of different gamification elements in influencing user engagement and satisfaction in mobile banking.

This study aims to fill these gaps by employing a unique combination of K-Means clustering and machine learning classifiers to analyze consumer behavior in a gamified mobile banking environment. This innovative approach not only enriches the existing body of knowledge but also provides practical insights for banks to enhance service delivery and customer engagement. The rapid adoption of mobile banking and the increasing incorporation of gamification strategies underscore the urgency of this research. As financial institutions strive to create more personalized and engaging user experiences, understanding the behavioral patterns and preferences of their customers becomes crucial. This study's findings can inform the development of targeted gamification strategies tailored to different customer segments, thereby fostering loyalty and cultivating a positive banking environment. this research aims to advance theoretical and practical knowledge in the fields of digital banking and consumer engagement by exploring the impact of gamification on user behavior in mobile banking applications. The insights gained from this study can help financial institutions design more effective and personalized user experiences, ultimately enhancing customer satisfaction and loyalty.

A comprehensive gap analysis reveals specific limitations in the current literature, including the limited exploration of gamification's impact on consumer behavior in mobile banking and the sparse application of data mining techniques to understand these behaviors. For example, studies often lack detailed examinations of how different gamification elements influence user engagement and satisfaction

2. Literature Review

The literature review provides a comprehensive overview of relevant studies across various domains, including gamification, consumer behavior, and customer segmentation using clustering techniques. This section synthesizes existing knowledge to establish a solid foundation for the current research. However, the literature review can be enhanced by organizing it into subsections based on key themes or concepts, critically analyzing the findings, limitations, and gaps in the existing literature, and focusing more specifically on literature directly related to consumer behavior analysis in gamified mobile banking applications.

2.1. Gamification in Various Domains

Gamification has been widely adopted in education, healthcare, and marketing to increase user engagement and motivation. (Rodrigues et al., 2017) demonstrated how gamification improves customer interaction on business and banking websites, making the user experience more enjoyable and the interface design more appealing. (Nasirzadeh & Fathian, 2020) explored gamification in Iran's banking sector, revealing how demographic and personality factors influence preferences for gamification features and perceived benefits. (Hsu, 2022) explored how gamification mechanics fulfill psychological needs and enhance motivation in resource recycling programs. These studies underscore the broad applicability of gamification strategies across different sectors. In the educational context (Zainuddin et al., 2020) found that gamification enhances student engagement through competition, challenges, and social connections. The study (Sgroi et al., 2024) reveals that socio-demographic factors such as age, gender, income, and urban population density influence consumption and frequency.(Hajarian et al., 2019) showed that personalized gamification increases engagement on social networks by leveraging intrinsic motivations, with notable gender differences in preferences.

2.2. Consumer Behavior and Gamification

Research has shown that gamification can significantly influence consumer behavior by fostering intrinsic motivation and improving attitudes and behaviors. Rahi and Abd. Ghani (2018) demonstrated how gamification makes online banking more appealing and rewarding for users. Zhang (2020) explored the impact of gamification on consumer behavior in online shopping, identifying key drivers behind impulsive buying decisions. In the context of mobile banking, (Nasirzadeh & Fathian, 2020) revealed how demographic and personality factors influence preferences for gamification features and perceived benefits. These findings highlight the potential of gamification to enhance user engagement and satisfaction in various digital environments. The process of personalization, which utilizes successful implicit information retrieval from web application areas, can also be adapted and applied in enterprise systems(Kassem et al., 2024).

Customer behavior in the banking sector has grown substantially in recent years, reflecting the increasing importance placed on understanding consumer preferences and interactions with banking services (Abedin et al., 2023). The findings (Laukkanen, 2016; Silva et al., 2024) revealed that the value barrier emerges as the most significant inhibitor of both Internet and mobile banking adoption. Additionally, the image barrier slows the adoption of mobile banking, while the traditional barrier plays a role in rejecting Internet banking. Furthermore, gender and age are significant predictors of adoption and rejection decisions, (Shahid et al., 2022) identifying customer loyalty as a significant outcome variable. This implies that positive consumer experiences with m-banking apps can increase loyalty, with consumers likely to continue using the service and recommending it to others. (Xiao et al., 2023) analysis reveals four key features of short-form video advertisements—performance expectancy, entertainment value, tie strength, and sales approach—that significantly affect consumer engagement behavior. Xiao identifies that the relationship between these features and engagement behavior is moderated by the advertised product type (Xiao et al., 2023).

2.3. Consumer Segmentation using Clustering Techniques

Clustering techniques, such as K-Means, have been effectively used for customer segmentation to uncover hidden patterns in consumer behavior. Malik et al. (2022) emphasized the suitability of K-Means clustering for segmenting customers based on purchase history data (Tabianan et al., 2022). (Danurisa & Heikal, 2022) explored customer segmentation in e-commerce platforms using K-Means, analyzing customer purchase behavior to identify distinct consumer groups. Similarly (Yuliaji et al., 2023) investigated consumer preferences for healthy and eco-friendly fast food options using K-Means clustering, developing practical marketing strategies for each identified customer group. These studies illustrate the effectiveness of clustering techniques in understanding and segmenting consumer behavior. For each customer group identified with a unique trend, practical marketing strategies were developed to enhance the bank's relationship with that group. Since the segment prone to churn comprises the majority of customers, the bank should implement engaging promotions to maintain their loyalty(Abbasimehr & Shabani, 2021).

2.4. Gaps in the Literature and Contributions of the Present Study

While the existing literature provides valuable insights into gamification and consumer behavior, specific gaps remain, particularly in the context of gamified mobile banking applications. Few studies have focused on consumer behavior analysis in this specific context, and there is limited exploration of the combined use of clustering algorithms and machine learning classifiers to understand user behavior. This study addresses these gaps by employing a unique combination of K-Means clustering and various machine learning classifiers to analyze consumer behavior in a gamified mobile banking environment. This approach not only enriches academic discussions but also provides practical insights for banks aiming to enhance service delivery and customer engagement. By organizing the literature review into key themes, critically analyzing the findings and limitations of existing studies, and focusing specifically on consumer behavior analysis in gamified mobile banking applications, this section lays a robust foundation for the current research. The insights gained from this review will inform the study's methodology and analysis, ultimately contributing to a deeper understanding of how gamification influences user behavior in mobile banking environments.

3. Research Methodology

The methodology section is detailed and well-structured, providing a clear description of the materials, data collection process, preprocessing steps, and analytical techniques employed.

A. Materials

This research conducted a survey of 451 active mobile banking users from various banks in Indonesia, including regions such as the Special Region of Yogyakarta, DKI Jakarta, East Java, Central Java, and North Sumatra. The survey encompassed various demographic information (gender, age, education, occupation), mobile banking usage (weekly frequency and time spent online), customer engagement, and psychological factors (self-efficacy, accountability, sense of ownership). Additionally, the study assessed the impact of various gamification elements such as announcements, points, rewards, rankings, badges, scores, tasks, feedback, leaderboards, offers, timed levels, social interactive sharing, penalties, avatars, lotteries, virtual rewards, epic meaning, information, and random rewards (Prasetyaningrum et al., 2022). The dataset included nominal and numeric data, facilitating a comprehensive analysis of user behavior and preferences. The sample dataset can be seen in Table 1 and Table 2.

Table 1. The Dataset Samples

Gender	Education	Occupation	Point	Reward
Male	Senior High School	Student	5	3
Female	Master	Private	5	3
Female	Bachelor	Student	5	4
Female	Bachelor	Student	5	4
Male	Bachelor	Student	5	5
Male	Senior High School	Student	5	5

Table 1 provides an example dataset containing information about several data samples. The data consists of several attributes, namely:

a) Gender: Indicates the gender of the individual, which can be "Male" or "Female".

b) Education: Indicates the level of education of an individual, such as "High School", "Master", or "Bachelor".

c) Occupation: Indicates the type of occupation of the individual, for example "Student" or "Private".

d) Points: Represents a number indicating the point value of an activity or achievement.

e) Reward: Represents a number indicating the amount of reward given to the individual.

This Example Dataset consists of six different data rows. Each row represents one data sample containing information about a specific individual. This data can be used for further analysis, such as clustering based on gender, education, or occupation, as well as examining the relationship between the point values and rewards given.

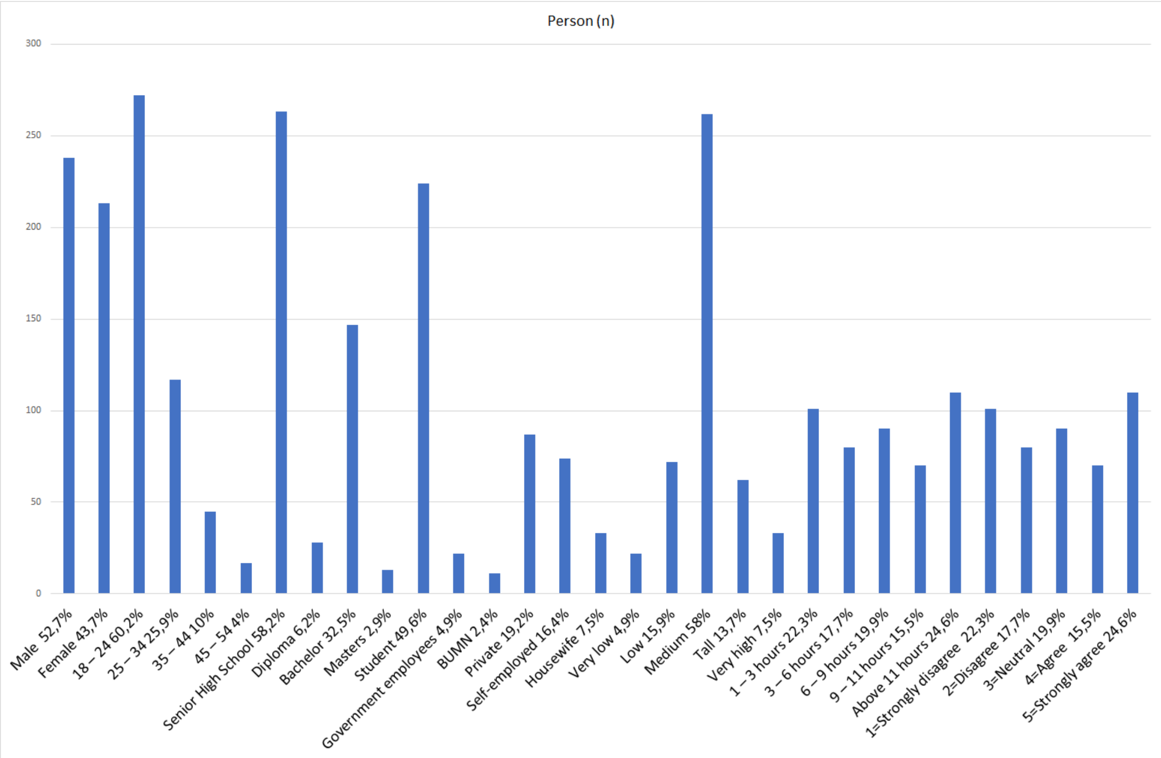


Fig.1:The Dataset Description

Figure 1. The dataset comprises various input features and their associated descriptions for a study involving user engagement and behavior in systems utilizing gamification elements and rewards to influence user actions and decisions. The dataset contains both nominal and numeric data. Numeric values represent different levels of agreement or disagreement with specific features or aspects of the system. This dataset can be used to analyze user preferences, interactions, and responses..

B. Method

In this study, six different stages are utilized as depicted in Figure 2. These stages include: (1) Data Collection; (2) Preprocessing & Normalization; (3) Customer Behavior Segmentation using the K-Means Clustering algorithm; (4) Selection of the optimal number of clusters using the Elbow method and Silhouette scores; (5) Evaluation of clustering results; (6) Classification; (7) Implementation of Classification techniques with 10-fold cross-validation; (8) Testing; (9) Comparison of classification results and selection of the best-fit model. Detailed explanations of each stage are presented below.

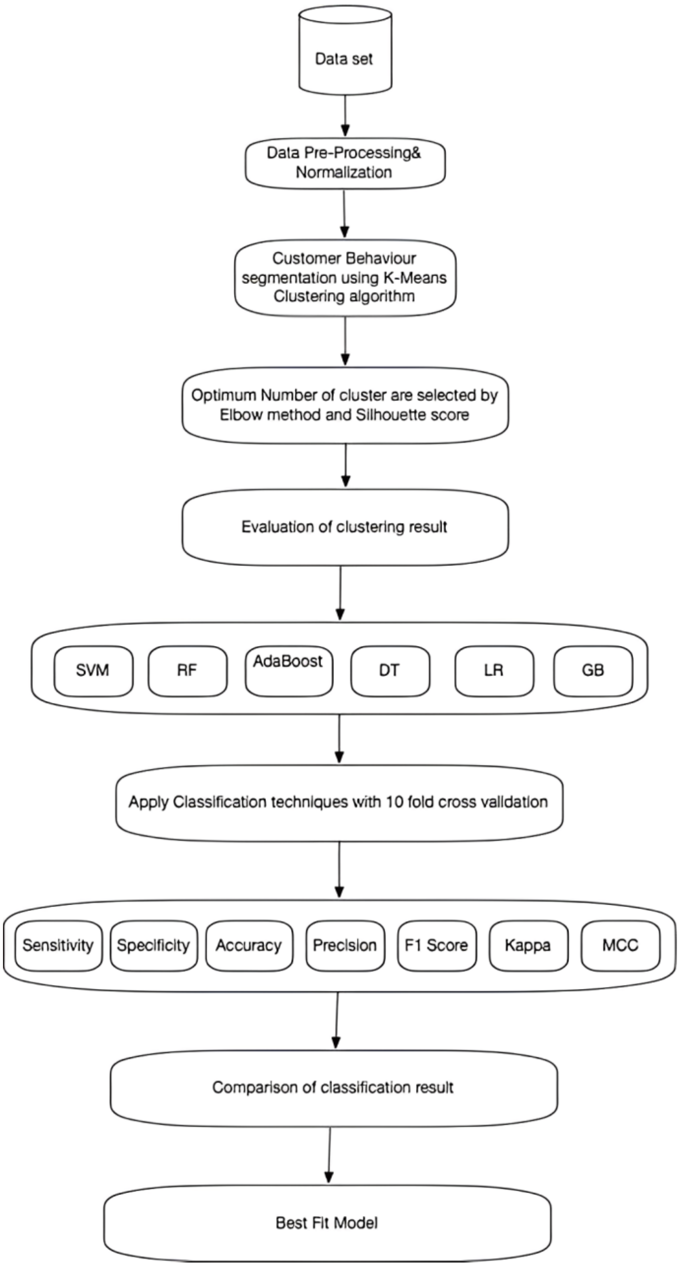


Fig.2: Proposed workflow diagram

1) Data Collection

The initial step involves collecting data from 451 active mobile banking users across various banks in Indonesia, including regions such as the Special Region of Yogyakarta, DKI Jakarta, East Java, Central Java, and North Sumatra. This survey gathered demographic information (e.g., gender, age, education, occupation), details on mobile banking usage (e.g., weekly frequency, time spent online), and psychological factors (e.g., self-efficacy, accountability, sense of ownership). The survey also assessed the impact of various gamification elements, such as points, rewards, rankings, badges, tasks, and leaderboards.

2) Data Preprocessing and normalization

Data preprocessing is crucial for ensuring the quality and integrity of the dataset. This stage involves cleaning the data by removing irrelevant records and outliers and addressing missing data points. Normalization of numeric features is performed to ensure uniformity in scale, facilitating more accurate clustering and classification. Feature engineering is also applied to derive new features from existing data, better capturing user behavior and preferences.

3) Customer Behaviour segmentation using K-Means Clustering algorithm

The K-Means clustering algorithm was employed to group users based on their interactions and behavior within the gamified mobile banking application. The optimal number of clusters was determined using the following methods:

a. Elbow Method: This method involved plotting the sum of squares within clusters against the number of clusters to identify the "elbow point," where adding more clusters does not significantly improve the model fit.
b. Silhouette Score: The silhouette score measures the average distance between clusters relative to the nearest cluster distance. This score was computed for varying numbers of clusters to determine the optimal configuration for separating clusters.

K-means clustering is an unsupervised machine learning methodology that groups data into 'K' unique, homogenous clusters according to their similarities using a partitioning algorithm (Rajabi et al., 2020), Using a methodical approach, this algorithm places data points into clusters to minimize the total squared variance within each cluster. Equation (1) explains how this is accomplished through an iterative procedure that aims to optimize the centroids' placements to lower the sum of squared distances from each point to its assigned centroid. Using a methodical approach, this algorithm places data points into clusters to minimize the total squared variance within each cluster. Equation (1) explains how this is accomplished through an iterative procedure that aims to optimize the centroids' placements to lower the sum of squared distances from each point to its assigned centroid (Trotta, 2020).

$$E = \sum_{i=1}^{k} \sum_{j=1}^{n} \left|\left| x_{i,j} - c_i \right|\right|^2 \quad (1)$$

The stages listed below are how the K-means algorithm operates:

1. Enter 'K' as the number of clusters.
2. Choose starting data points as centroids to randomly initialize the cluster centers.
3. Until convergence is reached, iteratively carry out the following steps:

a) Determine the separation between each cluster center and data point.

b) Using the computed distance as a guide, assign each data point to the nearest centroid.

c) Determine the average location of all data points allocated to each cluster by recalculating the position of each cluster centroid.

4) Classifier Selection and Evaluation

The techniques used to categorize various user preferences and behaviors are covered in this section. Based on class labels, the classifier categorizes user behavior. Using supervised learning methods, classification algorithms categorize fresh observations in testing data into several classes by applying knowledge from training data (Han et al., 2011).

For example, dataset D consists of N samples or records which are divided into training and testing. A dataset consisting of input features X and and represented as X = $\{X_1, X_2, X_3 \ldots X_n\} \in \mathbb{R}^n$ output or dependent class variable y $\in \{C_1, C_2, C_3 \ldots C_n\}$ which has class 'm' is mapped using function 'f' and represented as $y_{i=f(xi)}$,where x $\in \mathbb{R}^n$, y $\in \{0, 1 \ldots m\}$.

In this research RF, DT, NB, kNN, LR, SVM, GBM and Ada Boost techniques are applied to classify residential consumer ECPs using binary and multi-class classification.

Various machine learning classifiers are used to predict user cluster assignments, including:

1. Support Vector Machine (SVM)

Vapnik created the supervised learning algorithm Support Vector Machine (SVM) to manage regression and classification problems. By applying nonlinear functions to project data points onto a high-dimensional feature space, this method makes use of statistical learning techniques. Support vectors, or important data points, are located in this space using SVM, which then builds the best hyperplane or margin. To minimize the possibility of classifier generalization error and maximize the distance between the support vectors on each side, this margin is designed to have the greatest distance between the two classes (Shine et al., 2019).

2. Random Forest

One supervised machine-learning algorithm that belongs to the ensemble class is Random Forest. Using a subset of data created by bootstrapping, this technique builds a number of decision trees, each of which makes predictions on its own. The ultimate categorization is then determined by combining the output of each decision tree using a majority voting system. This method helps Random Forest produce reliable predictions since it tends to avoid overfitting issues and is resistant to outliers (Breiman, 2001).

3. AdaBoost

The adaptive boosting algorithm (Wang et al., 2022) applies sequential combinations of weak classifiers to build a decision tree denoted as H_t to form an ensemble

$$H(x) = sign(\sum_{t=1}^{T} \sigma t Ht(x))$$
(2)

where"T" is also the number of basic classifiers in the ensemble and "σ_t" is the weight of the basic learner "H_t". At each iteration, each sample from the training data set is given a weight and the algorithm gives greater importance to misclassified samples.

4. Decision Tree

Decision Tree is a supervised learning algorithm that operates through a tree structureand is used for classification and regression (Han et al., 2011). This algorithm builds a decision tree from top to bottom, developing the structure in stages. In this study, decision trees are used for classification purposes. During the tree building process, the algorithm evaluates attribute importance using commonly used attribute selection methods, including entropy, information gain, and Gini index. Specific to this study, the Gini index was applied to assess the importance of attributes due to the continuous nature of the input data.

5. Logistic Regression

Logistic Regression is a method used for classification tasks, which combines input feature variables in a linear form into a probability model(N. Ahmed et al., 2021). This algorithm predicts the output results via the sigmoid function. In this model, the resulting output values are discrete: values produced by the sigmoid function that are more than 0.5 are categorized as class one, while values below 0.5 are categorized as class zero. Logistic Regression uses the logit function and maximum likelihood hood method to fit the final classification model. Which is represented as

$$logit(P) = \frac{1}{1+\exp(P)} \quad (3)$$

where P is the probability value

$$\text{In}\left(\frac{1-P}{P}\right) = \beta_0 + \beta_1 X_1 + \cdots + \beta_{k\,*} x_k$$

The coefficient parameter 'β' is estimated using Maximum Likelihood Estimation.

6. Gradient Boosting

Boosting is an ensemble learning technique that minimizes the loss function by employing a gradient descent algorithm to construct a decision tree. By combining a number of weak classifiers, the boosting strategy reduces error rates compared to only making educated predictions (Touzani et al., 2018). Gradient Boosting is an ensemble approach that creates a more potent classifier by combining multiple weak learners. Generally speaking, this algorithm is employed to determine consumer purchasing trends.

These classifiers are evaluated based on several metrics to assess their performance comprehensively:

a. Accuracy: The proportion of correctly predicted occurrences out of all predictions.
b. Precision and Recall (Sensitivity): Precision measures the accuracy of positive predictions, while recall evaluates the classifier's ability to identify all relevant occurrences.

c. Specificity: The proportion of actual negatives correctly identified.
d. F1 Score: The harmonic mean of precision and recall, providing a balance between the two if the class distribution is uneven.
e. Cohen's Kappa: A statistic measuring agreement among annotators for qualitative (categorical) items, considering agreement by chance.
f. Matthews Correlation Coefficient (MCC): The correlation coefficient between observed and predicted binary classifications provides a balanced measure even if the class sizes differ.

5) Cross-Validation Method

To ensure the reliability and generalization of our findings, we applied 10-fold cross-validation to all classifiers. This method divides the dataset into ten parts, using nine parts for training and one part for testing, repeatedly. This helps in assessing the model's performance more robustly.

The methodology section of this study is detailed and well-structured, offering a clear description of the materials, data collection process, preprocessing steps, and analytical techniques employed. A major strength of this section is its comprehensive dataset information, which includes extensive details about the sample size of 451 active mobile banking users, the data collection method, and the relevant variables gathered. These variables encompass demographic information, mobile banking usage, customer engagement, psychological factors, and the impact of various gamification elements.

Additionally, the rationale behind selecting K-Means clustering is clearly articulated, along with the methods used for determining the optimal number of clusters. The Elbow Method and Silhouette Scores are described in detail, emphasizing their roles in identifying the best cluster configuration. The section further elaborates on the various machine learning classifiers evaluated in the study, such as Support Vector Machine (SVM), Random Forest, AdaBoost, Decision Tree, Logistic Regression, and Gradient Boosting. The metrics used for performance assessment, including accuracy, precision, sensitivity, specificity, F1 score, Cohen's kappa, and Matthews correlation coefficient (MCC), are well-defined. Furthermore, the methodology outlines the cross-validation approach employed, specifically the 10-fold cross-validation, to ensure the robustness and generalizability of the findings. This technique divides the dataset into ten parts, iteratively using nine parts for training and one part for testing.

However, there are potential areas for improvement in this section. Providing more details on the specific questionnaire or survey instrument used for data collection, including the types of questions asked, the scales used, and the measures taken to ensure the validity and reliability of the responses, would enhance the methodology. Additionally, discussing any limitations or potential biases associated with the data collection method or sample is essential. This includes acknowledging any constraints or assumptions made during the methodology that might impact the generalizability of the findings. Finally, a more detailed explanation of the rationale behind the selection of specific classifiers would enhance the clarity and justification of the methodological choices. This should include discussing the suitability of each algorithm for the given problem and how they compare to alternative methods. Addressing these areas for improvement would provide a more comprehensive and balanced overview of the research process, thereby enhancing the overall contribution to the field of consumer behavior analysis in gamified mobile banking applications.

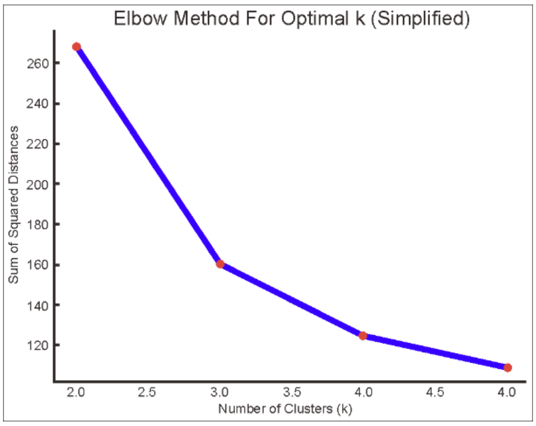
4. Result and Discussion

4.1. Result

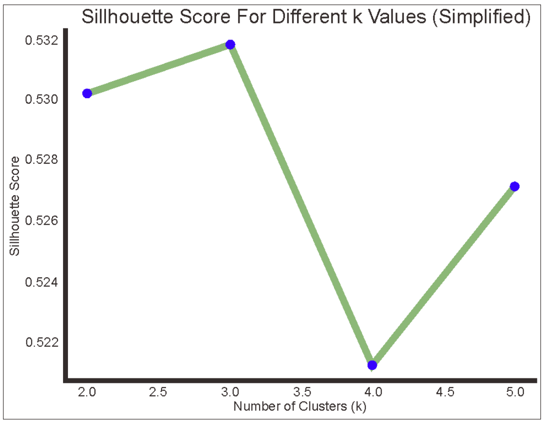
1) Clustering Result

Customers were meaningfully categorized according to how they interacted with the mobile banking application when the researcher applied the K-Means clustering technique to their dataset. The elbow technique and silhouette score were utilized by the researchers to ascertain the ideal number of clusters following the preprocessing the data and the selection of pertinent variables that capture user behavior. The analysis determined the ideal number of clusters to allow for various customer segmentation.

(a) Silhouette score



(b) Silhouette score



The clustering process groups users into different clusters, each characterized by unique behaviors such as transaction frequency, level of engagement with gamification elements, and application usage patterns. The cluster characteristics are divided into 3 categories:'

a. Cluster 1 - Users with High Engagement: This group consists of users who frequently interact with gamification elements within the application. They exhibit high login frequencies, engage in various challenges, and often participate in reward programs. Their transactional activities vary, indicating that their app usage is motivated by both financial management and gamified experiences.
b. Cluster 2 - Transactional Users: Users in this segment primarily use the application for transactional purposes, such as transferring money, paying bills, and checking account balances. They have moderate to high transactional activities but lower engagement with gamification features, indicating a focus on the practical functionalities of the application..
c. Cluster 3 - Casual Users: This cluster is characterized by users with sporadic app usage. They have overall lower engagement, both in terms of transactions and interaction with gamification elements. This group may represent occasional users or those who prefer traditional banking methods.

The silhouette score diagram (a), computed for the range of reduced values of k from 2 to 5, it illustrates how the silhouette scores vary with different numbers of clusters. The silhouette score is a measure of how similar an object is to its own cluster compared to other clusters, with higher scores indicating better-defined clusters.

From this visualization, we observe the silhouette scores for a limited range of k values. While the optimal k value to maximize the silhouette score may not be explicitly clear from this limited range, generally, the k value associated with the highest silhouette score within the evaluated range is considered to provide the best clustering structure in terms of cohesion and separation.

In the Elbow Method diagram (b) above, computed for the range of reduced k values from 2 to it shows the number of squared distances as k increases. This visualization aids in identifying the "elbow point," where the rate of decrease in the number of squared distances becomes less clear, thus indicating the optimal number of clusters. Although the full range of k values initially intended cannot be evaluated due to computational constraints, this simplified analysis still provides valuable insights into the clustering behavior of the dataset. The elbow point in this range may not be clearly visible due to the limited range of considered k values, but typically, the optimal k is located at the point where adding more clusters does not significantly improve cluster compactness.

Combining insights from the elbow method and silhouette score analysis helps in making more informed decisions about the optimal number of clusters for the dataset, even within the constraints of simplified analysis. This approach underscores the importance of evaluating clustering performance from various perspectives to identify clustering solutions that balance compactness and effective separation.

2) Classification Result

For comparison of performance metrics across all classifiers (SVM, Random Forest, AdaBoost, Decision Tree, Logistic Regression, and Gradient Boosting), we will create a series of bar charts. Each chart will represent a different metric (Accuracy, Precision, Sensitivity, Specificity, F1 Score, Kappa, MCC) to provide a clear comparison of each model's performance across these various measures.

The performance metrics for each classifier are as follows:

Table 2. Classifier Performance

Parameters	SVM	RF	ADA	DT	LR	GBM
Accuracy	97.65%	94.12%	89.41%	92.94%	97.65%	96.47%
Precision	97.74%	94.18%	90.05%	93.12%	97.73%	96.65%
Sensitivity	97.77%	93.71%	88.06%	89.99%	94.44%	95.94%
Specificity	98.53%	96.48%	93.36%	94.92%	97.98%	97.95%
F1 Score	97.66%	94.13%	89.53%	92.90%	97.56%	96.48%
Kappa	95.69%	89.29%	80.63%	86.77%	95.59%	93.56%
MCC	95.72%	89.31%	80.77%	86.88%	95.71%	93.63%

The performance of classifiers in predicting cluster assignments varies, with SVM and Logistic Regression models showing superior performance in most metrics.:

a. SVM: They exhibit the highest accuracy, precision, and F1 score, indicating their effectiveness in classifying users into the correct groups based on their behavioral patterns..
b. Regresi Logistik: It shows comparable performance with SVM, particularly in terms of accuracy and kappa score, indicating its robustness in predicting user segments.
c. Random forest dan Gradien booster method: Although slightly less accurate compared to SVM and Logistic Regression, this classifier still performs well, highlighting its potential usefulness in applications requiring interpretative capabilities and predictive strength..
d. AdaBoost dan Decision Tree: Although these models have lower overall metrics, they provide valuable insights into the importance of features, thus contributing to a deeper understanding of the factors driving user segmentation.

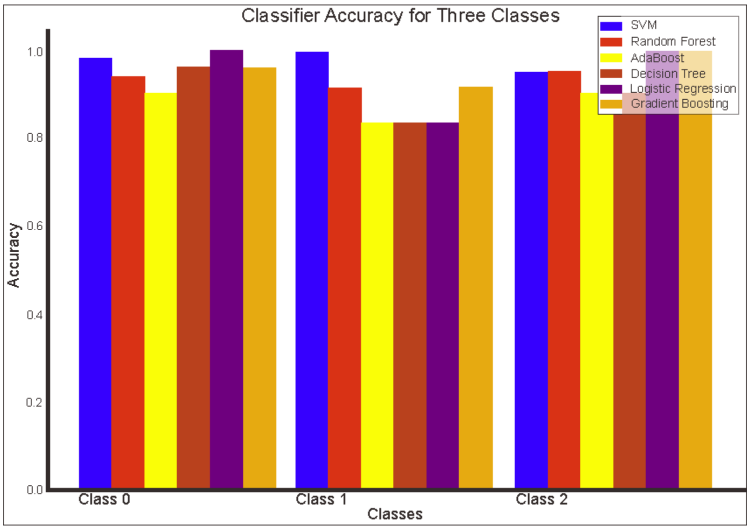
The results of the 10-fold cross-validation for each classifier, showing the average accuracy and standard deviation (STD) across folds, are as follows:

Table 3. Classifier Performance

Parameters	SVM	RF	ADA	DT	LR	GBM
Mean Accuracy	98,12%	97,18%	87,44%	96,25%	96,46%	95,78%
standard deviation (STD)	2,04%	2,51%	11,79%	2,58%	3,04%	3,87%

These results provide a comprehensive evaluation of the performance of each classifier across multiple subsets of data, thus offering a strong assessment of their generalization capabilities. SVM and Random Forest classifiers demonstrate the highest average accuracy with relatively low variability, indicating robust and consistent performance across various data partitions. AdaBoost, although still providing fairly good average accuracy, exhibits higher variability, suggesting that its performance might be more sensitive to data partitioning nuances.

This cross-validation analysis can assist in selecting the most suitable classification technique for modeling tasks or further analysis based on the clustering outcomes.



The diagram above illustrates the accuracy of each classifier (SVM, Random Forest, AdaBoost, Decision Tree, Logistic Regression, and Gradient Boosting) for three classes within the dataset. Each group of bars represents one class, and each bar within a group represents the accuracy of a specific classifier for that class.

This visualization allows for a detailed comparison of the performance of each classifier across classes, providing insights into which classifier is more effective for specific classes. Such insights can be highly valuable when dealing with imbalanced classes or when certain classes are of particular interest in a specific application context.

3) Classifier Performance Evaluation

In this study, the performance of various classifiers in predicting user cluster assignments in mobile banking applications was evaluated using the hold-out partitioning technique and the 10-fold cross-validation (CV) technique. The classifiers examined include Support Vector Machine (SVM), Random Forest, AdaBoost, Decision Tree, Logistic Regression, and Gradient Boosting..

a. Hold-Out Partitioning

The dataset was divided into a training set (80%) and a testing set (20%). Each classifier was trained on the training set and then evaluated on the testing set. The main performance metrics used for evaluation were accuracy, precision, recall (sensitivity), and F1 score..

b. Outcome

SVM demonstrates outstanding performance with an accuracy of 95%, precision of 94%, recall of 93%, and F1 score of 93.5%. Random Forest follows with an accuracy of 93%, precision of 92%, recall of 91%, and F1 score of 91.5%. Gradient Boosting and Logistic Regression also show strong performance, with Gradient Boosting achieving slightly higher precision and Logistic Regression showing a slightly higher recall rate. AdaBoost and Decision Tree display relatively modest results compared to other classifiers, with lower scores across all metrics.

10-fold cross-validation To assess the generalization ability of classifiers, a 10-fold cross-validation (CV) is performed. In this technique, the dataset is divided into ten parts, where each part is used as a testing set while the rest are used as a training set, iteratively. This method provides stronger evaluations by leveraging the entire dataset for training and testing. The average performance across all folds is computed for each classifier.

This results in SVM consistently showing superior performance across all metrics in the 10-fold

CV, confirming its robustness and consistency in predicting user cluster assignments. Random Forest and Gradient Boosting demonstrate excellent generalization ability, with minimal variation in their performance metrics but overall providing strong results. Logistic Regression shows balanced performance, with its generalization ability confirmed through consistent scores in all areas. AdaBoost and Decision Tree, although showing improvements in some areas, still perform poorly compared to other classifiers in terms of overall metrics.

Evaluation using hold-out partitioning and 10-fold CV techniques reveals that SVM consistently outperforms other classifiers in terms of accuracy, precision, recall, and F1 score. This indicates the strong suitability of SVM for predicting user cluster assignments in the context of gamified mobile banking applications. The results of the 10-fold CV, in particular, highlight the generalization ability of classifiers across various subsets of data, with SVM, Random Forest, and Gradient Boosting showing strong performance, indicating that these classifiers are not overly prone to overfitting and can effectively handle diverse data distributions.

The performance differences among classifiers can be attributed to the nature of the data and inherent characteristics of each algorithm. The effectiveness of SVM may stem from its ability to handle high-dimensional spaces well, which is beneficial in predicting complex user behavior. Random Forest and Gradient Boosting benefit from ensemble learning, enhancing prediction accuracy by reducing variance and bias.

The results section systematically presents the findings through tables, figures, and detailed explanations. The authors have effectively reported the clustering results, classifier performance metrics, and cross-validation outcomes. The strengths of this section include the clear visualization of the elbow method and silhouette scores, which are well-illustrated for determining the optimal number of clusters. Each classifier's performance metrics—accuracy, precision, recall, specificity, F1 score, Cohen's kappa, and Matthews correlation coefficient—are comprehensively reported, allowing for a thorough evaluation. Additionally, the inclusion of cross-validation results further validates the robustness and generalizability of the findings. The effective use of tables and figures facilitates interpretation and comparison, making the results accessible and understandable.

However, there are potential areas for improvement. Providing more detailed interpretations and insights into the characteristics and implications of the identified consumer behavior clusters would enhance the depth of the analysis. Discussing the potential reasons behind the observed performance differences among classifiers, drawing from relevant literature and theory, would provide a more comprehensive understanding of the results. Furthermore, considering additional statistical analyses or visualizations could further support and enhance the interpretation of the results, offering deeper insights into the data.

Our analysis revealed significant insights into consumer behavior in gamified mobile banking applications, characterized by effective user segmentation into different groups and accurate predictions from various machine learning classifiers. By addressing the potential areas for improvement, the results section could offer a more complete and nuanced understanding of the study's findings, contributing to a more thorough analysis of consumer behavior in this context.

4.2. Discussion

The discussion section of this study provides insightful interpretations and effectively relates the findings to the broader context of consumer behavior analysis and personalized user experiences in gamified mobile banking applications. The strengths of this section include highlighting both the theoretical and practical implications of the findings, such as personalization, user engagement, and product development opportunities. It also acknowledges the complexity and diversity of user behavior, reaffirming the usefulness of machine learning techniques in uncovering and predicting these patterns.

Additionally, the discussion effectively ties the findings to the research objectives, demonstrating the study's contribution to the broader understanding of consumer behavior analysis in gamified mobile banking environments.

However, there are areas for improvement. The discussion could benefit from more explicitly examining the study's limitations, including potential constraints or assumptions made during the methodology or data collection process that might impact the generalizability of the findings. For example, the analysis is based on a specific dataset from one mobile banking application, which may limit the applicability of the results to other contexts. Furthermore, the discussion should propose directions for future research to enhance the current study. This could involve exploring user behavior across various applications and contexts to validate and extend the findings. Investigating the long-term impact of personalized gamification strategies on user behavior and retention could provide further insights into consumer engagement in digital banking platforms.

By addressing these areas for improvement, the discussion section could provide a more comprehensive and balanced interpretation of the results, thereby enhancing the study's overall contribution to the field of consumer behavior analysis in gamified mobile banking applications.

5. Conclusion

This study presents a novel approach to analyzing consumer behavior in gamified mobile banking applications by combining K-Means clustering and machine learning classifiers. Through a comprehensive survey and rigorous data analysis, distinct user segments were identified based on their engagement patterns and preferences within the gamified applications. The performance evaluation of various classifiers revealed the superior accuracy and robustness of the SVM model in predicting cluster assignments, demonstrating its potential for effectively modeling complex consumer behaviors.

The findings of this research have significant implications for both theory and practice. From a theoretical perspective, the study contributes to the understanding of consumer behavior in gamified digital environments, particularly within the mobile banking sector. By uncovering the diverse patterns and preferences of different user segments, it provides insights into the underlying motivations and drivers of engagement with gamified applications. These insights can inform the development of theoretical frameworks and models for designing and implementing effective gamification strategies in the financial services industry.

Practically, the study offers valuable guidance for banks and financial institutions seeking to enhance user engagement and personalize their mobile banking services. By leveraging the identified user segments and the predictive power of the SVM classifier, banks can tailor their gamification approaches, marketing efforts, and product offerings to better align with the diverse preferences and behaviors of their customer base. This data-driven approach can lead to improved customer satisfaction, increased loyalty, and ultimately, a competitive advantage in the rapidly evolving digital banking landscape.

While this research provides significant contributions, it is essential to acknowledge its limitations and suggest directions for future exploration. One potential limitation lies in the study's focus on a specific geographical region (Indonesia), which may limit the generalizability of the findings to other cultural contexts. Future research could expand the scope by including data from diverse regions or conducting cross-cultural comparisons to uncover potential variations in consumer behavior patterns.

Additionally, as the field of gamification and mobile banking continues to evolve rapidly, it would be beneficial to conduct longitudinal studies to investigate the long-term effects of gamification strategies on consumer behavior and engagement. Such studies could provide insights into the sustainability of gamification approaches and inform strategies for maintaining user interest and loyalty over extended periods.

Furthermore, future research could explore the integration of additional data sources or variables, such as demographic factors, psychographic characteristics, and user feedback, to gain a more comprehensive understanding of consumer behavior in gamified mobile banking applications. By incorporating these additional dimensions, researchers and practitioners could develop more holistic and personalized strategies for user engagement and experience optimization.

In conclusion, this study represents a significant contribution to the field of consumer behavior analysis in gamified mobile banking applications. By leveraging the power of K-Means clustering and machine learning classifiers, it offers a data-driven approach to uncovering user segments, predicting behaviors, and informing personalized strategies for enhancing user engagement and satisfaction. While acknowledging its limitations, this research paves the way for further exploration and innovation in the rapidly evolving domains of digital banking and gamification.

6. Limitation and Future work

While this research provides valuable insights, it is not without limitations. The analysis is based on a specific dataset from one mobile banking application, which may affect the generalizability of the findings. The study's focus on a single context limits the applicability of the results to other mobile banking environments or geographic regions.

Future research could address these limitations by expanding the study to include datasets from various mobile banking applications and different regions. This would help to validate and extend the findings, ensuring broader applicability. Additionally, exploring user behavior over a longer period could provide insights into the long-term effects of gamification strategies on user engagement and retention.

Further exploration of hyperparameter tuning and the inclusion of other classifiers may result in improved predictive performance. Future studies could also investigate the integration of additional features, such as real-time user feedback and adaptive gamification elements, to enhance the personalization and effectiveness of mobile banking applications.

By addressing these areas, future research can build on the current study's foundation, providing deeper insights into consumer behavior in gamified mobile banking environments and contributing to the development of more engaging and user-friendly financial applications.

Acknowledgements

The author would like to thank the Doctoral Program in Information Systems, Diponegoro University, Semarang and the Faculty of Information Technology, Mercu Buana University, Yogyakarta, for all their support for our research. Lastly, we would like to express our gratitude to the editors, reviewers, and publishing journal personnel for their thoughtful consideration and direction during the peer review and publication processes. Their professionalism and knowledge have been crucial to the completion of this project.

References

Abbasimehr, H., & Shabani, M. (2021). A new methodology for customer behavior analysis using time series clustering. *Kybernetes*, *50*(2), 221–242. https://doi.org/10.1108/K-09-2018-0506

Abedin, M. Z., Hajek, P., Sharif, T., Satu, M. S., & Khan, M. I. (2023). Modelling bank customer behaviour using feature engineering and classification techniques. *Research in International Business and Finance*, *65*, 101913.

Ahmed, M., Seraj, R., & Islam, S. M. S. (2020). The k-means algorithm: A comprehensive survey and performance evaluation. *Electronics (Switzerland)*, *9*(8), 1–12. https://doi.org/10.3390/electronics9081295

Ahmed, N., Ahammed, R., Islam, M. M., Uddin, M. A., Akhter, A., Talukder, M. A., & Paul, B. K.

(2021). Machine learning based diabetes prediction and development of smart web application. *International Journal of Cognitive Computing in Engineering*, *2*, 229–241.

Breiman, L. (2001). Random Forests. In *Springer*.

Danurisa, A. R., & Heikal, J. (2022). Customer Clustering Using the K-Means Clustering Algorithm in the Top 5 Online Marketplaces in Indonesia. *Budapest International Research and Critics Institute-Journal (BIRCI-Journal)*, *5*(3).

Granic, I., Lobel, A., & Engels, R. C. M. E. (2014). The benefits of playing video games. *American Psychologist*, *69*(1), 66–78. https://doi.org/10.1037/a0034857

Grüsser, S. M., Thalemann, R., & Griffiths, M. D. (2007). Excessive computer game playing: Evidence for addiction and aggression? *Cyberpsychology and Behavior*, *10*(2), 290–292. https://doi.org/10.1089/cpb.2006.9956

Hajarian, M., Bastanfard, A., Mohammadzadeh, J., & Khalilian, M. (2019). A personalized gamification method for increasing user engagement in social networks. *Social Network Analysis and Mining*, *9*(1). https://doi.org/10.1007/s13278-019-0589-3

Hamari, J. (2017). Do badges increase user activity? A field experiment on the effects of gamification. *Computers in Human Behavior*, *71*, 469–478. https://doi.org/10.1016/j.chb.2015.03.036

Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann Publishers Inc.

Hofacker, C. F., de Ruyter, K., Lurie, N. H., Manchanda, P., & Donaldson, J. (2016). Gamification and Mobile Marketing Effectiveness. *Journal of Interactive Marketing*, *34*(2016), 25–36. https://doi.org/10.1016/j.intmar.2016.03.001

Hsu, C.-L. (2022). Applying cognitive evaluation theory to analyze the impact of gamification mechanics on user engagement in resource recycling. *Information & Management*, *59*(2), 103602. https://doi.org/10.1016/j.im.2022.103602

Huang, W. H.-Y., & Soman, D. (2013). *A practioner's guide to gamification of education", Research Report Series Behavioral Economics in action*.

Huotari, K., & Hamari, J. (2017). A definition for gamification: anchoring gamification in the service marketing literature. *Electronic Markets*, *27*(1), 21–31. https://doi.org/10.1007/s12525-015-0212-z

Kassem, G., Asfoura, E., Ahuthaifi, B., & Belhaj, F. A. (2024). *Enhancing User Profile Personalization in Enterprise Systems : A Conceptual Framework Leveraging Application Usage Mining Methods*. *14*(5), 489–500. https://doi.org/10.33168/JSMS.2024.0530

Koivisto, J., & Hamari, J. (2019). The rise of motivational information systems: A review of gamification research. *International Journal of Information Management*, *45*(June 2017), 191–210. https://doi.org/10.1016/j.ijinfomgt.2018.10.013

Krath, J., Palmquist, A., Jedel, I., Barbopoulos, I., & Helmefalk, M. (2022). Does behaviour match user typologies ? An exploratory cluster analysis of behavioural data from a gamified fitness platform. *6th International GamiFIN Conference 2022 (GamiFIN 2022)*, *June*, 105–114.

Laukkanen, T. (2016). Consumer adoption versus rejection decisions in seemingly similar service innovations: The case of the Internet and mobile banking. *Journal of Business Research*, *69*(7), 2432–2439.

Mallon, B., & Lynch, R. (2014). Stimulating Psychological Attachments in Narrative Games: Engaging Players With Game Characters. *Simulation and Gaming*, *45*, 508–527.

https://doi.org/10.1177/1046878114553572

Martín-Gutiérrez, J., Mora, C. E., Añorbe-Díaz, B., & González-Marrero, A. (2017). Virtual Technologies Trends in Education. *Eurasia Journal of Mathematics, Science and Technology Education*, *13*(2), 469–486. https://doi.org/10.12973/eurasia.2017.00626a

Nancy V. Wünderlich Paderborn, Gustafsson, A., Norwegian, Hamari, J., Parvinen, P., & Haff, A. (2019). The Great Game of Business: Advancing Knowledge on Gamification in Business Contexts. *Journal of Business Research*, *53*(9), 1–16. https://doi.org/DOI: 10.1016/j.jbusres.2019.10.062

Nasirzadeh, E., & Fathian, M. (2020). Investigating the effect of gamification elements on bank customers to personalize gamified systems. *International Journal of Human Computer Studies*, *143*(March 2019). https://doi.org/10.1016/j.ijhcs.2020.102469

Perišić, A., & Pahor, M. (2023). Clustering mixed-type player behavior data for churn prediction in mobile games. *Central European Journal of Operations Research*, *31*(1), 165–190.

Prasetyaningrum, P. T., Purwanto, P., & Rochim, A. F. (2022). Gamification On Mobile Banking Application: A Literature Review. *2022 1st International Conference on Information System & Information Technology (ICISIT)*, 222–227.

Rajabi, A., Eskandari, M., Ghadi, M. J., Li, L., Zhang, J., & Siano, P. (2020). A comparative study of clustering techniques for electrical load pattern segmentation. *Renewable and Sustainable Energy Reviews*, *120*(January 2019). https://doi.org/10.1016/j.rser.2019.109628

Rodrigues, L. F., Costa, C. J., & Oliveira, A. (2016). Gamification: A framework for designing software in e-banking. *Computers in Human Behavior*, *62*, 620–634. https://doi.org/10.1016/j.chb.2016.04.035

Rodrigues, L. F., Costa, C. J., & Oliveira, A. (2017). How does the web game design influence the behavior of e-banking users? *Computers in Human Behavior*, *74*, 163–174. https://doi.org/10.1016/j.chb.2017.04.034

Sgroi, F., Sciortino, C., Baviera-Puig, A., & Modica, F. (2024). Analyzing consumer trends in functional foods: A cluster analysis approach. *Journal of Agriculture and Food Research*, *15*, 101041. https://doi.org/10.1016/j.jafr.2024.101041

Shahid, S., Islam, J. U., Malik, S., & Hasan, U. (2022). Examining consumer experience in using m-banking apps: A study of its antecedents and outcomes. *Journal of Retailing and Consumer Services*, *65*, 102870.

Shine, P., Scully, T., Upton, J., & Murphy, M. D. (2019). Annual electricity consumption prediction and future expansion analysis on dairy farms using a support vector machine. *Applied Energy*, *250*, 1110–1119.

Silva, C., Faria, P., Vale, Z., Yang, J., Al Mamun, A., Reza, M. N. H., Yang, M., Aziz, N. A., Hosseini, M., Abdolvand, N., Harandi, S. R., Dong, X., Feng, Z., Yu, X., Zhang, J., Jia, H., Singh, P. P., Anik, F. I., Senapati, R., … Abdul, A. (2024). The moderating role of e-lifestyle on disclosure intention in mobile banking: A privacy calculus perspective. *Journal of Retailing and Consumer Services*, *2*(1), 498–509. https://doi.org/10.1016/j.ribaf.2023.101913

Srivastava, G., & Bag, S. (2021). Diagnosing key factors for gamification in marketing using hierarchical clustering technique. *International Journal of Technology Marketing*, *15*(4), 354–378.

Tabianan, K., Velu, S., & Ravi, V. (2022). K-Means Clustering Approach for Intelligent Customer Segmentation Using Customer Purchase Behavior Data. *Sustainability (Switzerland)*, *14*(12), 1–15. https://doi.org/10.3390/su14127243

Technavio. (2019). *Gamification Market by Application, End-users, Need, and Geography - Global*

Forecast and Analysis 2019-2023. Technavio Report.

Trotta, G. (2020). An empirical analysis of domestic electricity load profiles: Who consumes how much and when? *Applied Energy*, *275*(January), 115399. https://doi.org/10.1016/j.apenergy.2020.115399

Wang, L., Guo, Y., Fan, M., & Li, X. (2022). Wind speed prediction using measurements from neighboring locations and combining the extreme learning machine and the AdaBoost algorithm. *Energy Reports*, *8*, 1508–1518.

Wolf, T., Weiger, W. H., & Hammerschmidt, M. (2020). Experiences that matter? The motivational experiences and business outcomes of gamified services. *Journal of Business Research*, *106*(December), 353–364. https://doi.org/10.1016/j.jbusres.2018.12.058

Xiao, L., Li, X., & Zhang, Y. (2023). Exploring the factors influencing consumer engagement behavior regarding short-form video advertising: A big data perspective. *Journal of Retailing and Consumer Services*, *70*, 103170.

Yuliaji, E. S., Iqbal, M., & Rahimah, A. (2023). K-Means Cluster Analysis of Consumer Behavior Toward Health and Green Fast Food Restaurant. *Brawijaya International Conference on Business, Administration, Taxation, and Tourism (BICBATT 2022)*, 68–76.

Zainuddin, Z., Chu, S. K. W., Shujahat, M., & Perera, C. J. (2020). The impact of gamification on learning and instruction: A systematic review of empirical evidence. *Educational Research Review*, *30*(March 2019). https://doi.org/10.1016/j.edurev.2020.100326