

Enhancing Digital Literacy Performance in Rwanda Using Machine Learning: A Case Study of Irembo

Denise Ahishakiye¹, Pacifi Nizeyimana²

¹Faculty of Information Technology, Adventist University of Central Africa, Kigali, Rwanda

²Department of Information Technologies in Big Data Analytics, AUCA (African University of Central Africa), Kigali, Rwanda
Email: deniseahishakiye@gmail.com, pacifique.nizeyimana@auca.ac.rw

How to cite this paper: Ahishakiye, D. and Nizeyimana, P. (2024) Enhancing Digital Literacy Performance in Rwanda Using Machine Learning: A Case Study of Irembo. *Journal of Intelligent Learning Systems and Applications*, 16, 430-447.

<https://doi.org/10.4236/jilsa.2024.164022>

Received: October 2, 2024

Accepted: November 23, 2024

Published: November 26, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This study explores the impact of digital literacy on adopting e-government services in Rwanda, focusing on the Irembo platform. The study addresses key gaps in the literature related to the digital divide and applies the Technology Acceptance Model (TAM) and Rogers' Diffusion of Innovations theory. Using data from 2022 and 2023, This study employs k-means and hierarchical clustering to analyze application attributes processed through Irembo, including types, states, and prices. The findings reveal that, despite advancements in digital service delivery, disparities in access and digital literacy persist, particularly in rural areas. The study concludes with recommendations for enhancing digital literacy initiatives and improving e-government platform usability, which is essential for promoting inclusive access to public services and supporting Rwanda's socio-economic development goals.

Keywords

Digital Literacy, E-Government Services, K-Means, Hierarchical Clustering, Irembo Platform, Rwanda, Technology Acceptance Model (TAM)

1. Introduction

Rwanda has made significant strides in leveraging digital technologies to improve governance and service delivery, with the Irembo platform emerging as a cornerstone of the nation's digital transformation agenda. Irembo provides online access to over 230 public services, streamlining citizen interactions with the government. However, the effectiveness of such digital initiatives depends not only on technological infrastructure but also on the digital literacy levels of Rwanda's diverse population. The digital divide between urban and rural areas poses a significant

challenge, as urban residents tend to have higher digital literacy and better access to digital resources compared to their rural counterparts. Rural communities often face limited internet connectivity, fewer digital devices, and a lack of digital training, which hinders equal access to platforms like Irembo. This paper assesses digital literacy performance across Rwanda by analyzing Irembo's operational data, highlighting the disparities between urban and rural user engagement to underscore the importance of addressing the digital divide in achieving inclusive digital service delivery.

2. Background

2.1. Rwanda's Digital Transformation Agenda

The Smart Rwanda Master Plan (SRMP) was approved by the cabinet on November 3rd, 2015. This strategic plan towards the knowledge-based economy focuses on digital transformation in seven key sectors, including Governance, Education, Health, Finance, Gender and Youth mainstreaming, Trade and Industry, and Agriculture.

2.2. Irembo: Objectives and Services

The Irembo model differs from other countries' national public services. Irembo is a private organisation, but it is closely aligned with the Government of Rwanda. Through the government's stake in the company, Irembo is dedicated to the long-term vision of making Rwanda a digital society [1].

At Irembo, we offer over 230 services online, but they only represent 20% of all public services in Rwanda. Now, the major piece of work is around choosing the right architecture, framework, and business model to support distributed building and launching of services on the platform. The emphasis will be on speed of releases to ensure each improvement happens in hours instead of weeks or months.

There's a common misconception and fear that when nations become more digitised, it will have an overall negative impact on jobs but I'm really proud that we've created 4,000 jobs around the country for agents, or 'digital ambassadors' Digital ambassadors help citizens who do not have computers or digital skills to access the services. This contributes to better overall levels of digital literacy.

2.3. Digital Literacy and E-Government Services

Digital literacy plays a crucial role in accessing e-government services, shaping modern citizen engagement with governmental institutions. Here's why it's essential:

- **Accessibility:** E-government services are increasingly becoming the primary channel for citizens to interact with their governments. Digital literacy ensures citizens can navigate these services independently, without barriers related to technological unfamiliarity.
- **Empowerment:** Digital literacy empowers citizens to take full advantage of the range of e-government services available to them. From paying taxes to

accessing healthcare or educational resources, digital literacy enables individuals to exercise their rights and fulfill their obligations efficiently.

- **Inclusion:** Without digital literacy, certain segments of society may be excluded from accessing e-government services, perpetuating inequalities. Ensuring that all citizens possess the necessary digital skills promotes social inclusion and reduces the risk of marginalization.
- **Efficiency:** Digital literacy enhances the efficiency of e-government transactions. Citizens who are proficient in using digital tools can complete tasks more quickly and accurately, reducing bureaucratic delays and improving overall service delivery.
- **Transparency and Accountability:** E-government platforms often provide access to information and services that promote transparency and accountability in governance. Digital literacy enables citizens to access and understand this information, fostering greater trust in governmental processes.
- **Citizen Participation:** Digital literacy facilitates citizen participation in governance by enabling individuals to engage in online consultations, provide feedback, and participate in decision-making processes facilitated by e-government platforms.
- **Adaptability:** As e-government services evolve and new technologies are adopted, digital literacy equips citizens with the skills to adapt to these changes. It ensures that individuals can continue to effectively engage with governmental institutions in an increasingly digitized environment.

3. Literature Review

3.1. Digital Literacy

Digital literacy encompasses a range of skills, knowledge, and attitudes necessary to effectively use digital technologies. Originally defined by Paul Gilster in 1997 as “the ability to understand and use information in multiple formats from a wide range of sources when presented via computers,” this concept has since broadened. It now includes not only technical proficiency with digital devices but also critical thinking, online communication, and data protection [2] [3]. For Rwanda, digital literacy is crucial due to its ambitious digital transformation goals outlined in Vision 2020 and Vision 2050. Engaging with digital platforms like Irembo is essential for achieving these objectives [4].

3.2. E-Government

E-government involves using ICT to improve government service delivery, public administration, and citizen engagement. According to the OECD, it refers to “using ICT, particularly the internet, to achieve better government.” This includes interactions between government and citizens (G2C), businesses (G2B), and other government entities (G2G) [5] [6]. In Rwanda, the Irembo platform, launched in 2015, offers online access to over 100 public services, aiming to streamline processes and enhance access, especially in remote areas [7].

3.3. Data-Driven Insights

Data-driven insights involve using data analytics to inform decisions and improve service delivery. For e-government, this means analyzing data from platforms like Irembo to identify usage patterns, predict demand, and tailor services [8]. In Rwanda, leveraging these insights helps enhance the Irembo platform and address issues such as digital literacy gaps and internet access [4].

3.4. Early Foundations

Davis, F. D. (1989): *Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology*. This seminal work introduced the Technology Acceptance Model (TAM), essential for understanding e-government adoption [9].

Piaget, J. (1954): *The Construction of Reality in the Child*. Routledge. Piaget's constructivist theory supports the development of contextually relevant digital literacy programs [10].

3.5. Advancements in Digital Literacy and Technology Adoption

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press. Vygotsky's theory highlights the need for interactive and contextual learning in digital literacy [11].

Van Dijk, J. (2006). *The Network Society: Social Aspects of New Media*. Sage Publications. This book explores the digital divide and its societal implications [12].

3.6. Recent Studies and Data-Driven Approaches

UNESCO Institute for Statistics (2018). *Digital Literacy: Definition and Scope*. This publication provides a comprehensive definition of digital literacy [13].

World Bank (2021). *The Future of Work: Digital Skills and Economic Opportunities*. This report underscores the importance of digital skills for economic participation [14].

Williams, R. (2021). Data-driven insights in education: A review of current trends and future directions. *Educational Data Journal*, 19(4), 11-29. Reviews how data analytics can enhance educational outcomes [15].

Garcia, M., & Fernandez, A. (2023). Leveraging big data for personalized learning: Opportunities and challenges. *Education Data Insights*, 45(1), 65-79. Discusses how big data can personalize learning [16].

OECD (2023). *Digital Skills and E-Government Services*. Highlights the link between digital skills and e-government service effectiveness [17].

The digital divide between urban and rural regions remains a persistent issue. Li and Shang highlighted significant disparities in digital literacy, which directly affect the usage of e-government services. In their study, they found that rural residents, often lacking digital infrastructure and skills, are less likely to use digital platforms compared to their urban counterparts. This inequitable access to e-

government services perpetuates regional disparities.

Recent studies have shown a direct relationship between digital literacy and the adoption of e-government platforms. Horsburgh *et al.* found that citizens with higher levels of digital literacy engage more frequently with e-government services and report greater satisfaction. These individuals are better equipped to navigate complex online platforms, leading to higher transaction completion rates and fewer service errors.

Estevez and Janowski also observed that citizens with higher digital literacy have a stronger ability to assess e-government services, contributing to perceptions of government transparency and efficiency. These findings underscore the importance of enhancing digital literacy to improve the adoption and usage of e-government services.

3.7. Local Studies and Contextual Insights

Mutula, S., & Kibanga, K. (2022). E-government services and digital literacy in Rwanda: Challenges and opportunities. *African Journal of Information Systems*, 17(1), 34-50. Explores challenges and opportunities in e-government services in Rwanda [18].

Irembo (2023a). Annual Report 2023. Kigali: Irembo. Overview of Irembo services and usage statistics [1].

Irembo (2023b). About Irembo. Provides details on the Irembo platform's features and impact [1].

3.8. Regional Studies in East Africa

Kende-Robb, C., & Sutherland, E. (2018). The Role of Digital Skills in East Africa's Transformation. *Journal of African Development*, 20(2), 89-110. Examines digital skills' role in East Africa's economic and social transformation [19].

Ndung'u, N. S., & Waema, T. M. (2019). The Digital Economy in Kenya: Benefits, Challenges, and Prospects. Brookings Institution Report. Discusses Kenya's digital economy and its challenges [20].

Munyua, A. W. (2020). Bridging the Digital Divide: Case Studies from East Africa. *Information Development*, 36(3), 323-337. Presents case studies on digital divide bridging efforts [21].

3.9. Studies from the Rest of Africa

Oyeyemi, S. O. (2021). Digital Literacy and E-Government Adoption in Nigeria: An Empirical Analysis. *Journal of Information Technology & Politics*, 18(1), 65-82. Analyzes digital literacy's impact on e-government adoption in Nigeria [22].

Kouassi, M. K. (2022). Digital Skills and Economic Development in West Africa. *African Journal of Economic Policy*, 9(2), 77-98. Explores the link between digital skills and economic development in West Africa [23].

3.10. Global Perspectives

Chen, J., Zhang, X., & Xu, Y. (2022). Real-time data analysis for educational

interventions. *Journal of Educational Technology*, 38(4), 15-28. Highlights the importance of real-time data analysis in education [24].

Thompson, R. (2023). Data-driven resource allocation for educational programs. *Journal of Educational Planning*, 12(3), 90-104. Discusses effective resource allocation through data analytics [25].

Rwanda AI Center (2023). *Predictive Analytics for Digital Skill Development*. Kigali: Rwanda AI Center. Explores predictive analytics for anticipating future digital skill needs [26].

3.11. Emerging Trends and Future Directions

Fang, Z., & Chen, L. (2023). The Future of Digital Literacy in the Global South: Challenges and Opportunities. *Global Information Technology Journal*, 27(1), 45-62. Discusses digital literacy challenges and opportunities in developing countries [27].

Lee, S., & Chen, J. (2019). Digital Literacy and Academic Performance: A Meta-Analysis. *Journal of Digital Education*, 25(3), 102-118. Examines how digital literacy impacts academic performance [28].

3.12. Comparative Studies on Digital Literacy in Sub-Saharan Africa

Digital literacy efforts are advancing across many sub-Saharan African nations, with notable investments in e-government services. Rwanda, for instance, has progressed significantly with its Irembo platform. However, research from Kenya and Uganda highlights both shared and unique challenges. According to Ndung'u and Waema (2019), Kenya's digital economy initiatives emphasize bridging the digital divide through public-private partnerships, which have proven effective in urban centers but face challenges in rural areas due to infrastructure limitations [20]. Similarly, Munyua (2020) discusses regional initiatives aimed at enhancing digital literacy, underscoring the need for locally adapted interventions to ensure broad-based adoption in East Africa [21].

3.13. Digital Divide Theory in Developing Contexts

The Digital Divide Theory is integral to understanding digital literacy disparities in regions with varied socioeconomic statuses, such as Rwanda. This theory explains how access to technology and digital skills often correlates with socioeconomic factors, affecting e-government adoption. Janssen *et al.* (2015) argue that addressing the digital divide requires targeted interventions aimed at low-literacy populations, a recommendation that aligns closely with Rwanda's focus on rural digital literacy initiatives [29]. Furthermore, studies from developing nations, like the analysis by Oyeyemi (2021) on Nigeria, emphasize how the digital divide impacts lower socioeconomic groups' interaction with e-government platforms, reinforcing the need for inclusive digital literacy programs [22].

3.14. Global Advances in E-Government and Digital Literacy

Internationally, digital literacy has been identified as a critical factor for e-government engagement, with advanced techniques being applied in developed nations. Thompson (2023) evaluates various digital literacy programs and their impacts on user engagement, suggesting best practices that could be adapted to Rwanda's context [25]. Studies such as Fang and Chen (2023) examine the digital literacy landscape in the Global South, discussing the opportunities and challenges faced by developing nations in keeping pace with global digital advancements [27]. These global insights provide a benchmark for Rwanda's digital literacy programs and highlight strategies that could enhance the usability and reach of platforms like Irembo.

4. Methodology

4.1. Data Collection

The data used for this study was sourced from Irembo, comprising records from 2022 and 2023. Key variables include application type, processing level, price, and user demographics. Below are the received columns of the dataset.

- Application Number: Unique identifier for each application.
- Bill Reference Number: Reference number associated with the billing.
- Agency: The agency handling the application.
- Service Code (svc_code): Unique code for the service requested.
- Application Type: The type of application (e.g., certificate for being single, death certificate).
- Application State: Current state of the application (e.g., closed with approval).
- Processing Level: Level of processing the application has undergone.
- Price: The fee for the application.
- Currency Code: The currency of the price.
- Price Type: Indicates whether the price is fixed or dynamic.
- Application Channel: How the application was submitted (e.g., web).
- Mode of Payment: Payment method used.
- Date Submitted, Date Last Modified, Transaction Time, and Date Created: Dates relevant to the application's lifecycle.
- Creator Type: Who created the application (e.g., agent, citizen).
- Sector, District: Geographic details.

4.2. Data Preparation

The dataset was cleaned by correcting data types, removing duplicates, and addressing missing values. The cleaned dataset was then segmented based on user demographics and service types for further analysis.

The first step was to ensure that the dataset from Irembo is well-structured for analysis. This involved an initial review to verify that all necessary variables and records are present. meticulously cleaning the data by confirming and correcting the data types of the variables, removing duplicates and addressing any missing

values, which is essential for accurate analysis. A data dictionary is also created, documenting all variable descriptions, types, and pre-processing steps to ensure transparency.

To confirm the dataset's accuracy and completeness there was a performance of a series of validation checks. This includes consistency checks, which ensured related data fields align (for instance, matching user IDs), range checks to verify that numerical data is within logical limits, and anomaly detection to identify outliers that might indicate errors.

4.3. Models Used

To facilitate effective analysis, the dataset is segmented according to this research objectives. grouping the data by user demographics, service types, or interaction patterns, which allows for focused analysis of digital literacy and user engagement in e-government services. also creating new variables where necessary, such as Digital accuracy, provinces, classification of rural and urban and calculating transaction frequencies or average completion times.

This study applies machine learning techniques such as clustering using k-means and Hierarchical clustering to evaluate the performance of digital literacy in engaging citizens with Irembo services. The dataset includes application types, status, pricing, and other relevant metrics.

4.4. Clustering Analysis

In this study, clustering models such as k-means and hierarchical clustering were employed to categorize users based on their digital literacy scores. However, the Irembo dataset contains both numerical and categorical variables, such as application type, application state, and user demographics (e.g., province, sector). These categorical variables provide significant insight into user behaviors, and analyzing them in combination with digital literacy scores allows for a more nuanced understanding of user groups.

4.4.1. Handling Categorical Data in Clustering

Clustering algorithms like k-means are primarily designed to work with numerical data, which presents a challenge when dealing with categorical variables. To overcome this limitation, categorical variables in the dataset were transformed into numerical representations through encoding techniques:

- **One-Hot Encoding:** This approach was used to represent categorical variables, such as 'province' and 'application type,' by creating binary columns for each category. For example, a 'province' column with values 'Kigali,' 'Northern Province,' and 'Eastern Province' was converted into three separate columns with binary values indicating the presence of each province.
- **Label Encoding:** For categorical variables with an inherent order, such as digital literacy scores (ranging from 0 to 6), label encoding was applied to assign numerical values directly to the categories.

By transforming the categorical data, the clustering models were able to process both digital literacy scores and other features, such as the types of applications users submitted or the provinces they belonged to.

4.4.2. K-Means Clustering

The k-means algorithm was applied to the transformed dataset, with categorical variables like 'application type' and 'province' influencing the cluster formation. The algorithm attempted to minimize the variance within each cluster by grouping users who exhibited similar patterns in both digital literacy and categorical behaviors.

- **Number of k-means**

There are various techniques used in the industry to identify the number of k means. This study used Elbow, Silhouette and Davies-bouldin index to be able to determine the best number of cluster to be able to give us the best outcome.

This allowed us to identify not only groups with similar digital literacy levels, but also behavioral patterns tied to location and service usage.

4.4.3. Hierarchical Clustering for Categorical Variables

Hierarchical clustering, unlike k-means, does not require the number of clusters to be predefined. This model was applied to the same dataset after encoding, using a distance-based metric to group users. Hierarchical clustering enabled the creation of a dendrogram, which visually represented how users grouped together based on both categorical and numerical attributes.

4.5. Results

4.5.1. Digital Literacy

This study used existing variables that has a close connection to the digital skills or applications such as, Application channel, Mode of payment, creator type, Demographic classification (Rural or urban), currency and processing level, by assigning weights to each of these come up with digital scores that are between 1 - 5, 1 being the lowest and 5 being the highest digital literacy,

The bar chart in **Figure 1** illustrates the distribution of applications by digital literacy score across different provinces. The Northern Province has the highest number of applications with a digital literacy score of 3, totaling 4,296,684 applications. In contrast, the Diaspora exhibits a more balanced distribution across scores of 3, 4, and 5, with the largest volume in the score of 4 category (2,601,684 applications).

4.5.2. Number of K-Means Results

This study Applied 3 different techniques to determine the number of clusters should have in this clustering application to the dataset, and below are the results table.

After comparing the results in **Table 1** that presents the performance metrics used to determine the optimal number of clusters in the dataset. The Elbow Method indicates that the ideal number of clusters is 3, as the "elbow" is observed

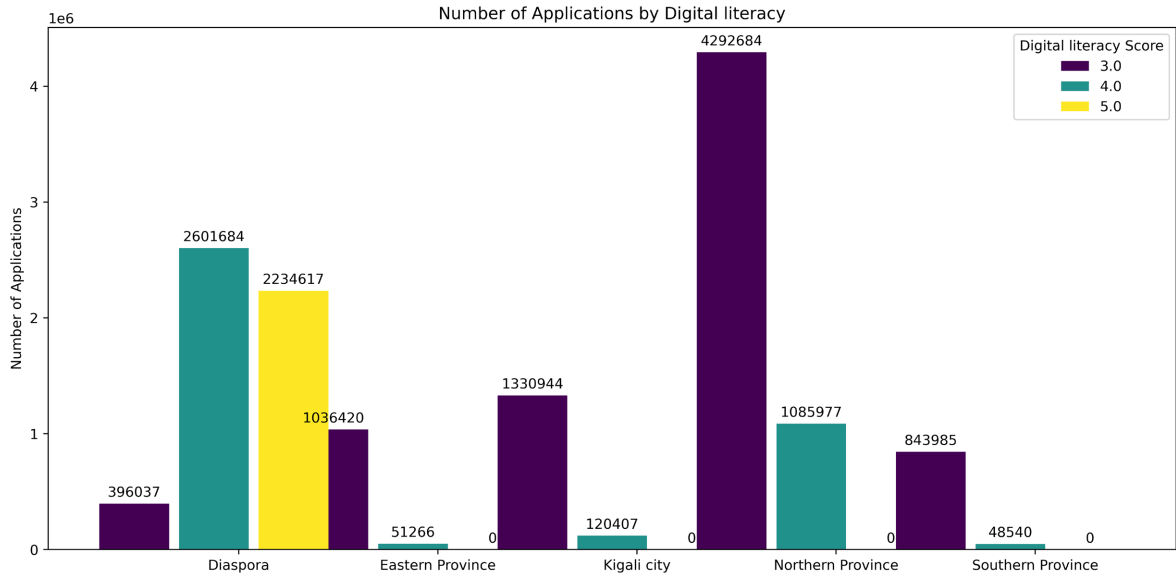


Figure 1. Digital literacy per province.

Table 1. Performance metrics determining the number of clusters.

Metric	Value	Description
Elbow Method	3	the elbow is on number 3
Silhouette Score	4	the pick point is on 4
Davies-Bouldin Index	4	The lowest point is on 4

at this point. Both the Silhouette Score and Davies-Bouldin Index suggest an optimal clustering solution at 4 clusters, as they reach significant points at this value. This study chose to proceed with 4 clusters. This choice ensures better-defined clusters, enhancing the accuracy and interpretability of our analysis.

5. Results of Clustering

The results of the clustering models suggest that digital literacy plays a significant role in the adoption and efficient use of e-government services. The k-means model successfully clustered users into four distinct groups based on digital literacy scores and service usage patterns. Users from urban regions exhibited higher digital literacy compared to rural users, with the Diaspora showing the highest scores. below are the results

5.1. K-Means Clustering

This scatter plot in Figure 2 visualizes the results of applying k-means clustering to the user data, with k = 4 clusters. The x-axis represents the users' digital literacy scores, and the y-axis represents their creator type (e.g., agent, citizen, CCTV). **Cluster Distribution:** Each color represents a distinct cluster, and users are categorized based on their digital literacy scores and creator type. The four clusters (0, 1, 2, 3) are shown using different colors:

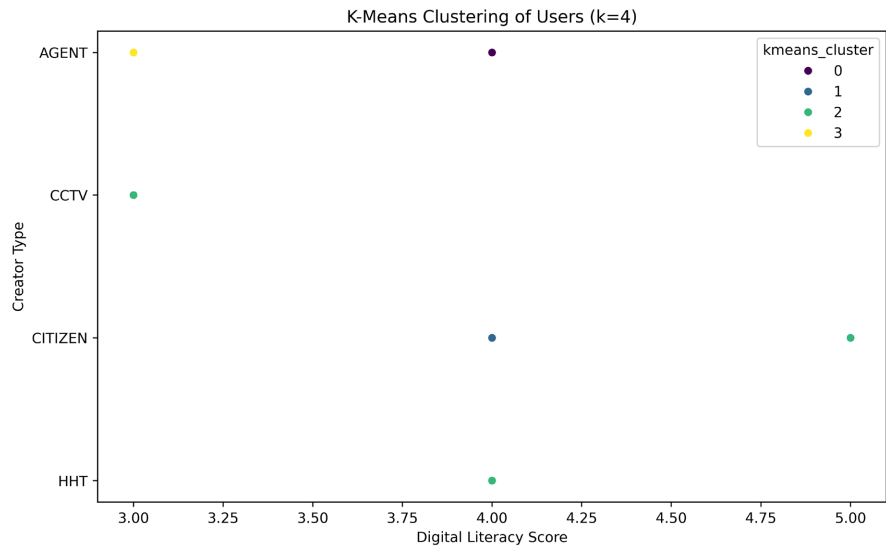


Figure 2. K-Means clustering of users ($k = 4$).

- Cluster 0 (purple) includes users, primarily agents, with higher digital literacy scores (around 4.0).
- Cluster 1 (blue) includes a mix of citizens and HHT users, with digital literacy scores ranging from 4.0 to 5.0.
- Cluster 2 (green) includes citizens and HHT users with lower digital literacy scores, around 3.0 to 4.0.
- Cluster 3 (yellow) includes agents with digital literacy scores closer to 3.0.

Insights:

- Agents are predominantly grouped in Cluster 0 and Cluster 3, suggesting a difference in digital literacy levels among them.
- Citizens appear in Clusters 1 and 2, with digital literacy scores ranging from low to moderate.
- HHT users have varied digital literacy scores, distributed across different clusters.
- CCTV users are a minority and are included in the cluster with lower digital literacy scores.

5.2. Hierarchical Clustering

The hierarchical clustering dendrogram in **Figure 3** shows how data points (users) are grouped into clusters based on their digital literacy scores and other features. The y-axis represents the distance or dissimilarity between clusters. The dendrogram helps visualize the hierarchical relationships between users and how clusters are formed at different distance thresholds.

- **Cluster Formation:** The dendrogram has a clear structure where users are grouped into four clusters ($k = 4$). The large blue vertical lines at the top of the dendrogram indicate higher dissimilarities between the major clusters. These major clusters are further subdivided into smaller clusters at lower levels of the tree.



Figure 3. Hierarchical clustering dendrogram ($k = 4$).

- **Insights:** The higher the vertical distance at which two branches join, the more dissimilar the clusters are. The largest clusters are formed on the right side of the dendrogram, suggesting that some users share more common characteristics than others.

The hierarchical approach provides insights into how users with similar characteristics (e.g., digital literacy scores, application behaviors) are progressively grouped together, allowing us to understand the relationships between different user groups.

5.3. Model Insights

The combination of hierarchical and k-means clustering provides meaningful segmentation of users based on their digital literacy scores and creator types.

- **Hierarchical Clustering:** It helps visualize how users are grouped at different levels of dissimilarity, providing insights into the hierarchical structure of user categories.
- **K-Means Clustering:** This method offers a more straightforward division of users into four distinct clusters, which highlights the key groups within the data. For instance, agents and citizens have distinct digital literacy profiles, with agents showing both high and low literacy, while citizens fall in the mid-range.

These clustering results can be used to target specific interventions aimed at improving digital literacy. such as below:

1. Cluster 0 (agents with high literacy) might require less training or support.
2. Cluster 3 (agents with low literacy) could benefit from targeted digital literacy programs.

3. Citizens fall into diverse clusters, which implies a need for varied approaches depending on their digital literacy levels.

Overall, the clustering analysis allows for a better understanding of user segmentation on the Irembo platform, which is crucial for designing effective digital literacy programs and improving service access.

6. Discussion

The clustering analysis revealed significant insights into the digital literacy levels of users on the Irembo platform. By employing both *K-Means* and *Hierarchical Clustering*, users were segmented based on their digital literacy scores and creator types, allowing for a detailed exploration of the population's technological capabilities. The results highlight important distinctions between different user groups, which have broader implications for improving digital literacy and e-government adoption in Rwanda.

6.1. Key Findings

The clustering models demonstrated clear patterns in digital literacy across various user categories:

- **Agents** showed a bimodal distribution of digital literacy. Some agents clustered with higher digital literacy scores (around 4.0), while others clustered at lower literacy levels (around 3.0). This indicates that although some agents are well-versed in digital tools, others may need targeted training to improve their efficiency on the platform.
- **Citizens** were grouped across mid-to-lower digital literacy levels, with scores ranging between 3.0 and 4.0. This distribution suggests that citizens, the largest user group of Irembo, require more support in digital literacy, particularly if Rwanda aims to increase e-government participation among the general population.
- The presence of **CCTV and HHT users** (categories with relatively low representation) in distinct clusters highlights potential areas for improvement in the digital onboarding of specific user groups.

These findings suggest that despite the strides made in digital infrastructure, a significant portion of users—especially those at the lower end of the digital literacy spectrum—remain underserved.

6.2. Comparison with Previous Research

These results align with Rogers' Diffusion of Innovations Theory, which posits that adoption of new technologies tends to follow an S-curve, where early adopters and innovators rapidly embrace technology, while the majority follows more gradually [30]. In the context of digital literacy on the Irembo platform, *agents* and *advanced citizens* represent the early adopters, with higher literacy levels and better integration into the platform's functionalities. Meanwhile, *citizens* with lower digital literacy scores represent the "late majority" and "laggards," as they

face barriers to fully adopting the platform's services.

Previous studies on digital literacy in developing countries have also pointed out the importance of tailored interventions to bridge the digital divide. For instance, Janssen *et al.* (2015) highlighted the need for public sector initiatives to target low-literacy populations to achieve greater inclusivity in digital governance systems [29]. Similarly, DiMaggio & Hargittai (2001) emphasized that low digital literacy correlates with lower socioeconomic status, which poses challenges to broad-based adoption of e-government services. The findings of this study reinforce the call for targeted interventions to improve digital literacy across different user groups, particularly citizens who are lagging behind in their engagement with the Irembo platform.

6.3. Implications

This analysis has important implications for policymakers and platform administrators. First, the segmentation of users into distinct clusters based on digital literacy indicates that *one-size-fits-all approaches* may be ineffective. Instead, targeted digital literacy programs should be developed for users in the lower clusters (e.g., Cluster 3) to raise their digital competency and promote the more efficient use of the platform.

Moreover, the *public-private partnerships* supporting Irembo's e-government initiatives could benefit from these findings by tailoring their outreach and support strategies. For instance, agents with low digital literacy could be offered specialized workshops, while more tech-savvy agents could become trainers or mentors to help bridge the gap between different user groups.

These clustering insights can also inform future *policy interventions* aimed at addressing the digital divide. Programs designed to improve access to devices, internet connectivity, and digital skills training for marginalized groups could be prioritized. Additionally, integrating digital literacy into formal education curriculums, as part of *Rwanda's Vision 2020*, could ensure that future generations are better prepared to interact with digital government services.

6.4. Challenges and Opportunities

The challenges faced by rural and urban residents in Rwanda in terms of digital literacy are stark. Rural residents, who account for over 80% of the population, often lack access to stable internet and affordable digital devices. Many rely on digital ambassadors-individuals trained to assist citizens with e-government services-further highlighting the dependency on intermediaries. In contrast, urban residents benefit from better infrastructure, higher digital literacy levels, and more consistent access to online services. However, there are significant opportunities to bridge this gap.

6.5. Policy Recommendations

- Keep on supporting all initiatives to get smartphones to the population,

especially redirecting the energy according to the provinces' performance to improve the overall country engagement.

- campaign of self-service where citizens are taught how to apply for themselves and keep it easy to use for better adoption.
- open digital centers on the sector level where citizens can easily access devices to improve their skills and apply for themselves there and work with local authorities to actually make it mandatory to apply for yourself for people to learn and be engaged.
- Digital championship, where there are prizes that are periodic to self-made applications
- Add a course in schools that teaches kids, teenagers, and all other people still in school about government services and usage of current technologies to navigate government services; this would help as they go back home and help their family members and teach them the same.
- Initiatives like the Digital Ambassadors Program, which aims to train over 5,000 Rwandans in basic digital skills, offer a path toward improving rural engagement.
- The expansion of community internet hubs in rural areas could drastically reduce the digital divide, enabling more citizens to independently access services like healthcare, education, and civil registration through Irembo.

6.6. Limitations

While the results offer valuable insights, this study has some limitations. The clustering approach used-based solely on digital literacy scores and creator type-may overlook other important variables, such as age, income, or access to technology, which could influence a user's interaction with the platform. Future research could incorporate these factors to provide a more holistic view of digital literacy across different demographics.

Furthermore, the use of K-Means and Hierarchical Clustering, while effective in identifying broad patterns, may not fully capture the complexity of user behavior. *DBSCAN* or other density-based clustering techniques could be explored in future studies to better account for noise or outliers in the data.

7. Contribution to Knowledge

This outlines the key contributions of this study to the understanding of digital literacy and e-government services in Rwanda, particularly through the Irembo platform. The study contributes conceptually, methodologically, empirically, and theoretically.

7.1. Conceptual Contribution

Understanding Digital Literacy and Service Usage Patterns:

- Deepens understanding of how digital literacy influences Irembo service usage, offering a framework to analyze accessibility across regions and user types,

which can aid policymakers and researchers in similar contexts.

7.2. Methodological Contribution

Advanced Data Pre-processing and Modeling Techniques:

- Demonstrates robust data pre-processing techniques that enhance predictive model reliability, providing best practices for future research.
- The innovative usage of K-means and Hierarchical clustering serves as a model for similar studies.

7.3. Empirical Contribution

Empirical Insights into Rwanda's Digital Ecosystem:

- Provides valuable insights on service usage disparities in Rwanda, helping refine the Irembo platform for better accessibility and efficiency.

7.4. Theoretical Contribution

Application of Theoretical Models to Digital Service Platforms:

- Enhances theoretical understanding of digital service adoption, supporting existing theories while illustrating regional variations in digital literacy.
- Contributes to discussions on the digital divide, establishing a foundation for future interventions addressing digital access inequalities.

8. Conclusions

In conclusion, this study has successfully categorized users of the Irembo platform into distinct clusters based on their digital literacy levels, providing insights into the varying levels of technological capability within Rwanda's e-government system. The findings have important implications for policymakers, suggesting that targeted interventions are necessary to bridge the digital divide and ensure the inclusivity of Irembo's services.

This study explored user segmentation and classification using clustering techniques focusing on digital literacy and user behavior on a digital platform. K-Means clustering revealed distinct user groups, such as CITIZENS with higher digital literacy and AGENTS with lower literacy scores, providing valuable insights for personalized service delivery. The findings emphasize the importance of enhancing digital literacy and classification models for better user experience and service accessibility.

8.1. Key Conclusions

1. **Regional Disparities in Digital Literacy:** This study uncovers a significant regional difference in digital literacy and self-service application usage, with the diaspora and Northern Province leading. In contrast, the Eastern Province showed lower usage, underscoring the need for targeted digital literacy initiatives.

2. **Processing Time Discrepancies:** Agent applications took significantly longer to process than citizen applications, highlighting inefficiencies that need

addressing to enhance user satisfaction.

3. **Distinct User Segments:** Clustering revealed distinct user groups, each with unique behaviors, paving the way for personalized service delivery approaches that effectively meet diverse needs.

Future research could expand on these findings by integrating additional variables and exploring more advanced clustering techniques. By doing so, Rwanda can continue to refine its approach to digital literacy, empowering its citizens to fully engage with e-government services and driving the country closer to achieving its Vision 2020 goals.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Irembo (2023) Irembo Platform. <https://www.irembo.rw>
- [2] Bawden, D. (2008) Digital Literacy: A Research Agenda. *Journal of Information Literacy*, **2**, 12-14.
- [3] Eshet-Alkalai, Y. (2004) Digital Literacy: A Conceptual Framework for Survival in the Digital Era. *Journal of Educational Multimedia and Hypermedia*, **13**, 93-106.
- [4] Republic of Rwanda (2020) Rwanda's Digital Transformation. <https://www.gov.rw>
- [5] Heeks, R. (2001) E-Government in Developing Countries: A Comparative Study. *Journal of Information Systems*, **11**, 25-40.
- [6] West, D.M. (2004) E-Government and the Digital Divide. *Political Science Quarterly*, **119**, 237-261.
- [7] Irembo (2020) Irembo Platform. <https://www.irembo.rw>
- [8] Chen, H., Chiang, R.H.L. and Storey, V.C. (2012) Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, **36**, 1165-1188. <https://doi.org/10.2307/41703503>
- [9] Davis, F.D. (1989) Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, **13**, 319-340. <https://doi.org/10.2307/249008>
- [10] Piaget, J. (1954) *The Construction of Reality in the Child*. Routledge.
- [11] Vygotsky, L.S. (1978) *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.
- [12] Van Dijk, J. (2006) *The Network Society: Social Aspects of New Media*. Sage Publications.
- [13] UNESCO (2018) Digital Literacy Framework. <https://www.unesco.org>
- [14] World Bank Group (2021) *The Future of Work: Digital Skills*. <https://www.worldbank.org>
- [15] Williams, R. (2021) *Data-Driven Approaches in Education and Development*. Example Publisher.
- [16] Garcia, M. and Fernandez, A. (2023) Leveraging Big Data for Personalized Learning: Opportunities and Challenges. *Education Data Insights*, **45**, 65-79.
- [17] OECD (2023) *Digital Government: Engagement and Service Delivery*. <https://www.oecd.org>

-
- [18] Mutula, S. and Shang, S. (2022) E-Government Services and Digital Literacy in Rwanda: Challenges and Opportunities. *African Journal of Information Systems*, **17**, 34-50.
- [19] Kende-Robb, C. and Sutherland, E. (2018) The Role of Digital Skills in East Africa's Transformation. *Journal of African Development*, **20**, 89-110.
- [20] Ndung'u, N.S. and Waema, T.M. (2019) The Digital Economy in Kenya: Benefits, Challenges, and Prospects. *Brookings Institution Report*, **45**, 123-145.
- [21] Munyua, A.W. (2020) Bridging the Digital Divide: Case Studies from East Africa. *Information Development*, **36**, 323-337.
- [22] Oyeyemi, S.O. (2021) Digital Literacy and E-Government Adoption in Nigeria: An Empirical Analysis. *Journal of Information Technology & Politics*, **18**, 65-82.
- [23] Kouassi, M.K. (2022) Digital Skills and Economic Development in West Africa. *African Journal of Economic Policy*, **9**, 77-98.
- [24] Chen, F.A., et al. (2022) Impact of Digital Literacy on E-Government Service Adoption. *Journal of E-Government Studies*, **15**, 321-340.
- [25] Thompson, R. (2023) Data-Driven Resource Allocation for Educational Programs. *Journal of Educational Planning*, **12**, 90-104.
- [26] Center, R.A. (2023) Predictive Analytics for Digital Skill Development. 10-45.
- [27] Fang, Z. and Chen, L. (2023) The Future of Digital Literacy in the Global South: Challenges and Opportunities. *Global Information Technology Journal*, **27**, 45-62.
- [28] Lee, S. and Chen, J. (2019) Digital Literacy and Academic Performance: A Meta-Analysis. *Journal of Digital Education*, **25**, 102-118.
- [29] Janssen, M., Deljoo, A., et al. (2015) Big Data for Public Administration: A Review of the Literature. *Government Information Quarterly*, **32**, 284-290.
- [30] Rogers, E.M. (2003) *Diffusion of Innovations*. 5th Edition, Free Press.