**Title: Interpretation of Cluster Drivers for Financial Inclusion in Nigeria Using SHAP Values and KMeans Clustering**

**Introduction**

This report discusses the clustering of individuals in Nigeria with respect to financial inclusion using KMeans clustering and SHAP (SHapley Additive exPlanations) values. In this analysis, we assess how various features affect cluster distribution. This can help us understand the characteristics of the variables that influences financial inclusion and economic opportunity among populations.

**Objective** To analyze the features with the most impact for the cluster assisgnment for financial inclusion dataset using KMeans model and SHAP values to explain the result.

**Methodology**

* **KMeans Clustering Algorithm**: KMeans was used to group populations into clusters.
* **Explainable AI(SHAP)**: SHAP values were employed to explore the feature(s) with the most impact to the cluster distribution.
* **Visualization**: The bar plot were used to interpret and visualize the feature importance.

**Results and Interpretation**

The SHAP summary bar plot reveals the average impact of each feature on the formation of three clusters (Cluster 0, Cluster 1, Cluster 2):

|  |  |  |
| --- | --- | --- |
| **Feature** | **Cluster Impact** | **Interpretation** |
| **Age** | All (esp. 0 & 1) | Dominant feature influencing clusters. Separate clusters as follows:  Cluster 0: middle-aged Individuals (Average Age: 36).  Cluster 1: Young Individuals (Average Age: 22).  Cluster 2: Older individuals (Average Age: 57). |
| **fin30 (Utility payment)** | 0 & 1 | Shows financial transactions across individuals |
| **Education (education level)** | 0 & 1 | Educational level influences financial behavior patterns. |
| **Employment Status (emp\_in)** | All | Plays a moderate role, indicates employment status across individuals |
| **Account Ownership (account)** | 0 | Helps identify individuals with formal bank accounts. |
| **Internet Access** | All | Indicate digitally connected individuals. |
| **Received Agricultural Payments** | All | Indicates financial activity in rural/agrarian sectors. |
| **Savings Behavior (saved)** | All | Small but relevant for distinguishing savers. |
| **Mobile Ownership** | All | Slightly affects clustering, aligned with digital finance access and services. |
| **Borrowed Money** | All | Shows individual credit behavior. |
| **Income Quartile** | All | Shows the income range of individuals and how it affects their financial habits. |

Other features such as female, receive\_wages, and urbanicity\_f2f had less impact but were included within the clusters.

**5. Implications for Financial Inclusion and Economic Opportunity**

* **Age** is a major determinant, indicating differing financial behaviors across generations.
* **Education and financial knowledge** significantly affect inclusion, suggesting education initiatives could reduce exclusion.
* **Account ownership**, although not the top factor, still shapes clusters, confirming its relevance.
* **Technology access (internet, mobile)** correlates with financial inclusion, implying digital finance is a key pathway.
* **Region** the rural or urban had less impact on the financial inclusion.
* **Utility Payment** contributes significantly to financial transactions across individuals.

**Conclusion**

SHAP explanations offer valuable insights into the features affecting financial inclusion. This analysis shows that while demographic features such as age dominate, financial behaviors and access to technology and education level also made significant contributions. This can help Policymakers and financial institutions to better design initiatives that will help reduce financial inclusion gaps.

**Recommendations**

* Create financial education awareness programs across age groups.
* Employ mobile/digital platforms for access to financial services.
* Monitor account and credit usage for segmentation.