Report: Automating Detection of Forged Banknotes

Prepared for: *First Bank of Nigeria*
Prepared by: *OGUNTOLA ABIODUN*
Date: *22-06-2025*
Project Title: *Banknote Vision — Clustering the Unseen Truth*

1. Purpose of the Project

The goal of this project was to explore whether machine learning, specifically clustering algorithms, could help identify forged banknotes based on digital characteristics. If successful, this could pave the way for automating detection and reducing manual inspection in the bank's workflow

2. Description of the Data

The dataset used was the **Banknote Authentication Dataset**, which includes over a thousand samples of scanned banknote images. For each note, several numeric features were recorded based on wavelet-transformed images:

- **Variance**
- **Skewness**
- **Kurtosis**
- **Entropy**

These features reflect the internal texture and shading patterns of the banknotes, patterns that can vary between genuine and fake notes, often in subtle ways invisible to the human eye.

3. Methods and Analysis

The approach consisted of the following steps:

- Exploratory analysis to study how the features are distributed, including summary statistics (mean, standard deviation) and visualizations (histograms, scatter plots).
- Data preprocessing: All features were standardized to ensure fair comparisons (scaling them to have equal weight in the analysis).

- **PCA (Principal Component Analysis)** was used to reduce the complexity of the data into 2 dimensions, allowing us to visualize patterns more clearly.

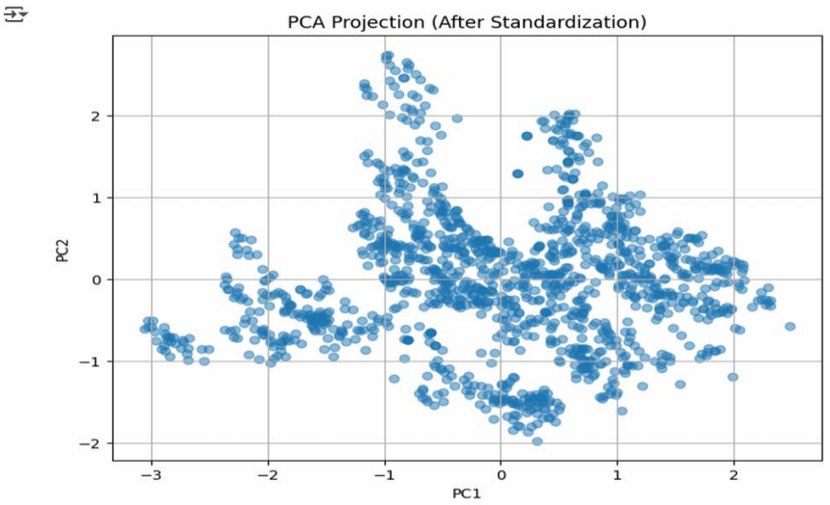

Figure 1

- K-Means clustering was applied to group banknotes into two or more clusters based solely on their characteristics, without using the true “genuine vs. forged” labels.
- Multiple runs of K-Means were performed to assess stability and consistency.

4. Summary of Results

- The algorithm consistently grouped banknotes into two distinct clusters, which aligned well with the real-world categories of **genuine** and **forged**.

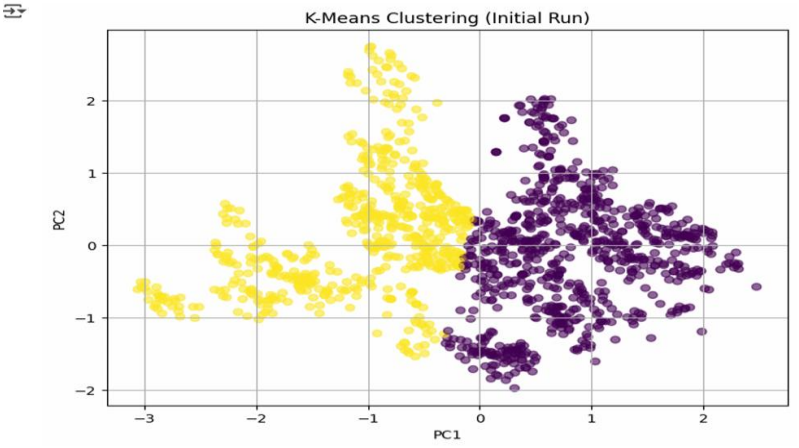

Figure 2

- The clustering appeared stable across multiple runs, with very few notes changing groups. This shows reliability in the model's ability to separate different types of banknotes based solely on patterns in the data.

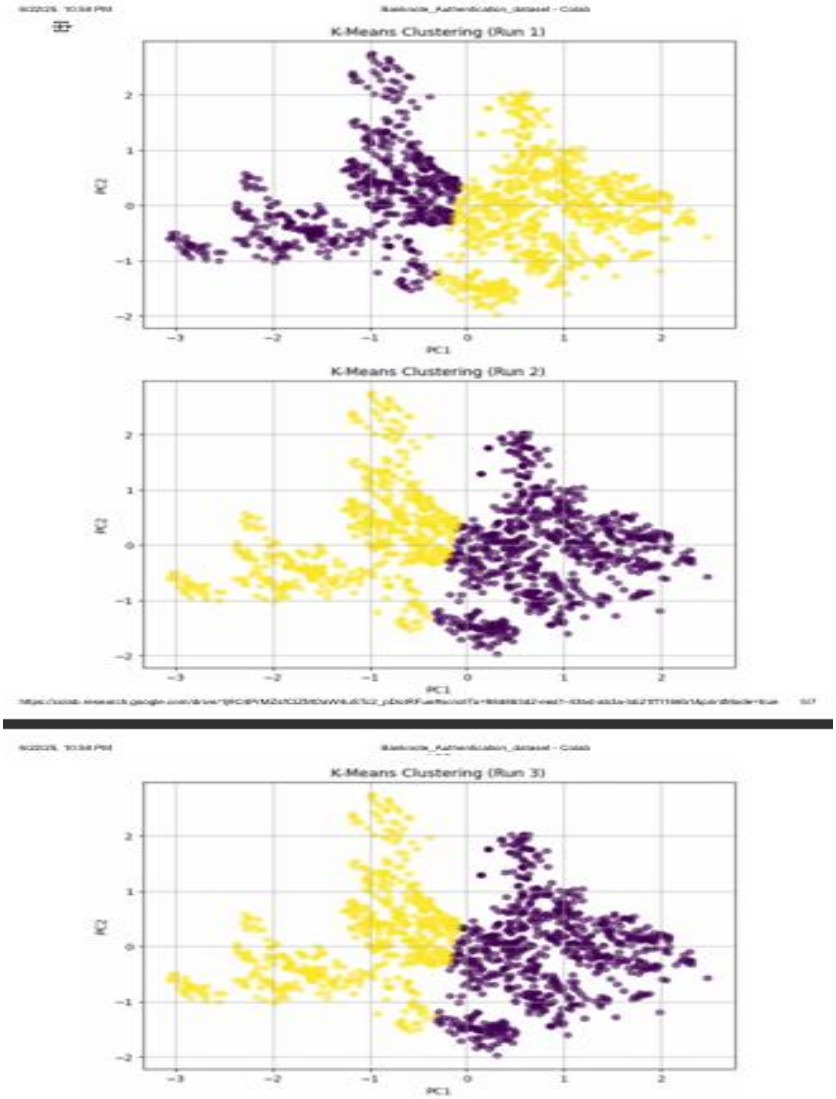

Figure 3

- Visualizations such as the PCA scatter plots reinforced that there is real structure in the data that machine learning can detect.

5. Recommendations

Based on these findings, we recommend the following:

- ✅ **Pilot a machine learning system** for automated screening of banknotes using similar digital feature extraction and clustering models. This will support human inspection or help prioritize which notes need review.
- ✅ **Collect more data**, especially modern or higher-resolution samples, to continue training the model and improve accuracy over time.
- ✅ **Combine clustering with supervised classification** in future iterations, allowing the model not only to group but to confidently predict “genuine” or “forged” labels with high accuracy.
- ✅ **Integrate this technology** into cash-handling systems or ATMs where real-time classification could prevent fraud attempts

Closing Remark

Clustering models like K-Means offer a powerful first step toward automating forged note detection. While not flawless on their own, they provide a strong backbone for AI-assisted decision-making and highlight where further attention is needed. This project confirms that the banknote dataset carries meaningful signals that machines can learn from, and puts your institution one step ahead in securing financial transactions