# AccessAI: Detecting Anomalous Swipe Behavior in Physical Access Control Systems

Divyaansh Vats
Data Science Intern
Fidrox Technologies Pvt Ltd

June 24, 2025

## 1 Project Overview

As a Data Science Intern at Fidrox Technologies Pvt Ltd, I contributed to the AccessAI project, which aims to detect anomalous swipe behavior in physical access control systems using unsupervised learning. My work focused on developing a simulated dataset and implementing K-Means clustering to identify unusual access patterns that could indicate security risks, such as unauthorized access or badge misuse. The project lays the groundwork for enhancing security in environments like corporate offices or data centers by analyzing swipe card data.

## 2 Work Performed

I created a Jupyter Notebook (`FIDROX_K_MEANS.ipynb`) to generate synthetic swipe data and perform initial clustering analysis. Key tasks included:

- **Data Generation**: Developed a dataset of 31,505 swipe records over 30 days, simulating 1,000 swipes daily across 100 users (U001–U100) and eight doors (e.g., Main Entrance, Server Room). Features included:
  - **Timestamp**: 80% of swipes during business hours (8 AM–7 PM), 20% during off-hours to mimic anomalies.
  - **DoorID and DoorName**: Weighted access frequencies (e.g., 25% for Main Entrance, 5% for Server Room).
  - **Direction**: Randomly assigned "IN" or "OUT".
  - **Result**: 95% "Success", 5% "Failed", with 1% of records featuring clusters of rapid failed attempts to simulate suspicious behavior.
- **Data Preprocessing**: Normalized features using `StandardScaler` and engineered time-based features (e.g., hour, day) for clustering.
- **K-Means Clustering**: Tuned the number of clusters using the Elbow Method and Silhouette Score, fitted the model, and visualized clusters via PCA in 2D

scatter plots.

- **Anomaly Detection**: Defined an anomaly score based on distance from cluster centroids, flagging the top 5% farthest points as potential anomalies.
- **Visualizations**: Created PCA scatter plots, bar charts of failed swipes per door, and heatmaps of failed swipes by hour and door.
- **Cluster Analysis**: Summarized cluster size, failed swipes, and anomaly counts, visualized via bar charts comparing failure and anomaly rates.
- **Data Storage**: Saved the dataset as `AccessAI_Simulated_Swipe_Logs.csv` for future use.
- **Tools**: Used Python libraries (`pandas`, `numpy`, `scikit-learn`, `matplotlib`, `seaborn`) on a GPU-accelerated Google Colab environment (T4 GPU).

# 3 Technical Approach

The synthetic dataset was designed to reflect realistic access patterns, with anomalies introduced via off-hour swipes and clusters of failed attempts. K-Means clustering grouped similar swipe behaviors, and outliers were identified based on their distance from cluster centroids. PCA enabled visualization of high-dimensional data, while the Elbow Method and Silhouette Score ensured optimal cluster selection. The approach highlighted failed swipes and odd-hour accesses as key indicators of suspicious behavior, aligning with real-world security concerns.

# 4 Key Observations

K-Means outliers frequently corresponded to failed swipe attempts or off-hour accesses, indicating that distance-based anomalies capture behaviorally suspicious patterns. Compared to DBSCAN (also explored), K-Means excelled at categorizing user behaviors into coherent groups, while DBSCAN was better for isolating sparse anomalies. This dual approach suggests K-Means is ideal for behavioral analytics and DBSCAN for pure anomaly detection.

# 5 Future Steps

Future work includes implementing additional algorithms (e.g., Isolation Forest, One-Class SVM), enhancing features (e.g., UserRole, DayOfWeek), and developing interactive dashboards with Streamlit or Dash for real-time monitoring. Validating models with real-world data will further ensure practical applicability.

# 6 Conclusion

My contributions to AccessAI established a robust foundation for detecting anomalous swipe behavior using K-Means clustering. By generating a realistic dataset and implementing clustering techniques, I demonstrated the potential to identify security risks in access control systems, supporting Fidrox Technologies' mission to deliver innovative, data-driven security solutions.