

Predictive Analysis of Aadhaar Biometric Authentication Failures

Karan Sharma¹, Adit Hajre¹, and Soham Sandip Mhatre¹

^aUIDAI Data Hackathon 2026

Problem Statement and Approach

Aadhaar biometric authentication serves as the primary mode of identity verification for accessing welfare schemes and essential public services. In practice, biometric authentication may fail due to worn fingerprints, aging-related degradation, sensor quality issues, or adverse operating environments. Such failures often force citizens to rely on demographic authentication or lead to repeated retries, increasing friction and the risk of service denial.

Despite the operational importance of biometric failures, public UIDAI datasets do not explicitly label authentication attempts as failed. This limitation motivates a predictive approach that infers biometric failure risk using indirect but observable authentication behavior. Specifically, this work models biometric stress through three dimensions: (i) loss of biometric dominance, (ii) increased reliance on demographic fallback authentication, and (iii) elevated authentication intensity caused by repeated attempts. These dimensions are combined into a composite risk score computed at the level of date, state, district, and pincode, enabling UIDAI to proactively identify high-risk regions and populations and intervene before service denial occurs.

Datasets Used

The analysis exclusively uses Aadhaar datasets provided by UIDAI, ensuring full compliance with hackathon requirements and data governance norms.

The **Biometric Authentication Dataset** provides age-wise counts of successful biometric authentications, which serve as the basis for measuring biometric dominance and coverage.

The **Demographic Authentication Dataset** records authentication events completed using demographic attributes when biometric authentication is not used, serving as a direct observable proxy for biometric stress and fallback dependency.

The **Enrolment Dataset** contains age-wise enrolment counts, enabling normalization of authentication activity and construction of population-relative indicators such as authentication intensity.

All datasets are aggregated by date, state, district, and pincode. Age-group alignment is performed to ensure that enrolment and authentication metrics correspond to the same population segments.

Methodology

This section details the complete analytical pipeline used to infer biometric authentication failure risk in the absence of explicit failure labels.

Data Integration and Aggregation. The biometric, demographic, and enrolment datasets were merged using common

geographic and temporal keys. Due to the transactional nature of authentication logs, multiple records may exist for the same location and date. These records were aggregated using summation to obtain total authentication volumes and enrolment counts per unit of analysis.

The analysis focuses on the 18+ age group, which represents the majority of Aadhaar-based service usage and exhibits higher biometric variability due to aging and occupational factors.

Core Metric Construction

Let B_{18+} denote biometric authentications, D_{18+} demographic (fallback) authentications, and E_{18+} the enrolled population for a given unit.

Total authentication demand is defined as:

$$T_{18+} = B_{18+} + D_{18+}$$

Biometric dominance is captured through biometric share:

$$\text{Biometric Share} = \frac{B_{18+}}{T_{18+}}$$

Fallback dependency is measured as:

$$\text{Fallback Dependency Ratio} = \frac{D_{18+}}{T_{18+}}$$

Authentication intensity, capturing retry behavior and friction, is defined as:

$$\text{Authentication Intensity} = \frac{T_{18+}}{E_{18+}}$$

Authentication intensity is normalized to ensure comparability across regions.

Significance Statement

Biometric authentication failures within the Aadhaar ecosystem can result in denial of essential public services, particularly for elderly citizens, manual laborers, and populations with degraded biometric quality. As public UIDAI datasets do not explicitly record biometric failure events, this study proposes a predictive, risk-based framework that infers biometric stress using indirect but observable authentication indicators. By combining biometric dominance, demographic fallback dependency, and authentication intensity into a composite risk score, the framework enables UIDAI to proactively identify high-risk regions and populations and deploy alternate authentication mechanisms such as OTP or Iris before service denial occurs.

The team jointly conducted data integration, feature engineering, and biometric risk modeling. All team members contributed to problem formulation, interpretation of analytical results, and preparation of the final report.

The authors declare no conflict of interest.

Composite Risk Score and Categorization. A weakly supervised biometric failure risk score is constructed as a weighted combination of interpretable authentication stress indicators:

$$\text{Risk Score} = 0.5 \times \text{Fallback Dependency Ratio} + 0.3 \times (1 - \text{Biometric Share}) + 0.2 \times \text{Normalized Authentication Intensity}$$

The resulting continuous risk score is discretized into three operationally meaningful categories to support decision-making by UIDAI:

- **Low Risk:** Risk Score < 0.3
- **Medium Risk:** $0.3 \leq \text{Risk Score} < 0.6$
- **High Risk:** Risk Score ≥ 0.6

Each risk category is directly mapped to a recommended UIDAI operational response, ensuring that analytical outputs translate into actionable interventions.

Risk Level	Recommended UIDAI Action
Low	Normal biometric authentication flow
Medium	Prepare and prioritize fallback authentication mechanisms
High	Proactively enable OTP or Iris authentication to prevent service denial

Data Analysis and Visualisation

This section presents key analytical findings derived from the constructed biometric failure risk indicators. The visualisations illustrate how fallback dependency, biometric dominance, authentication intensity, and regional variation collectively explain biometric stress patterns across Aadhaar authentication systems.

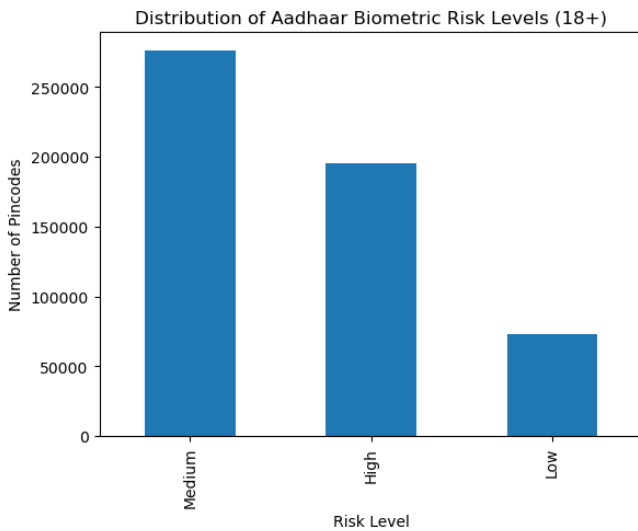


Fig. 1. Distribution of Aadhaar Biometric Failure Risk Levels (18+). A significant proportion of pincodes fall under medium and high risk categories, highlighting the need for proactive authentication planning.

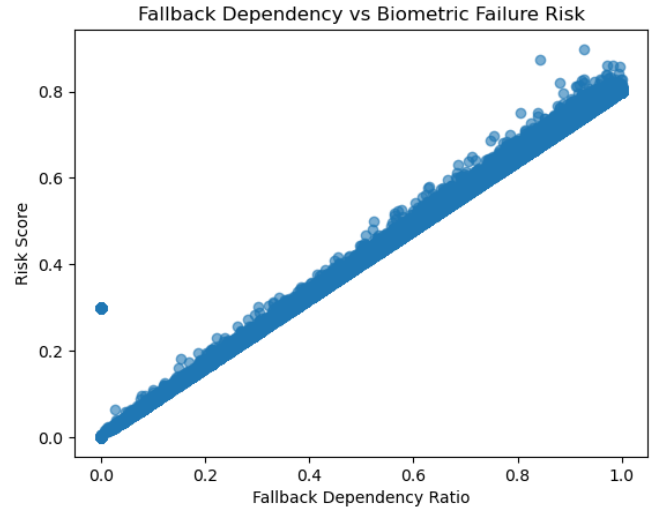


Fig. 2. Relationship between fallback dependency ratio and biometric failure risk score. Higher reliance on demographic authentication strongly correlates with increased biometric stress, validating fallback usage as a primary failure proxy.

The combined evidence from these visualisations demonstrates that biometric failure risk is not uniformly distributed and is strongly influenced by fallback dependency, biometric dominance, and authentication friction. These insights support data-driven, region-specific policy and operational decisions by UIDAI.

All analyses and visualisations were conducted using Python, with complete code and derived metrics provided in the accompanying Jupyter Notebook.

The combined evidence from these visualisations demonstrates that biometric failure risk is not uniformly distributed and is strongly influenced by fallback dependency, biometric dominance, and authentication friction. These insights support data-driven, region-specific policy and operational decisions by UIDAI.

All analyses and visualisations were conducted using Python, with complete code and derived metrics provided in the accompanying Jupyter Notebook.

Code Availability

The complete source code, data preprocessing scripts, and analysis notebooks used in this study are publicly available at the following repository:

<https://github.com/cooldued2004/UIDAI>

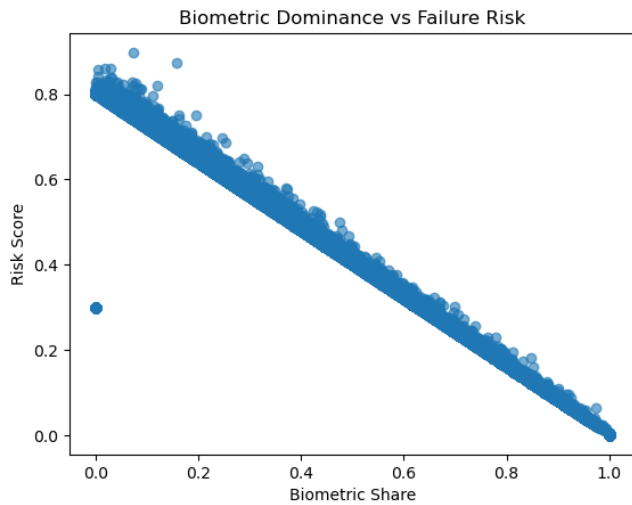


Fig. 3. Inverse relationship between biometric share and biometric failure risk. Regions with higher biometric dominance exhibit lower failure risk, emphasizing the protective role of robust biometric coverage.

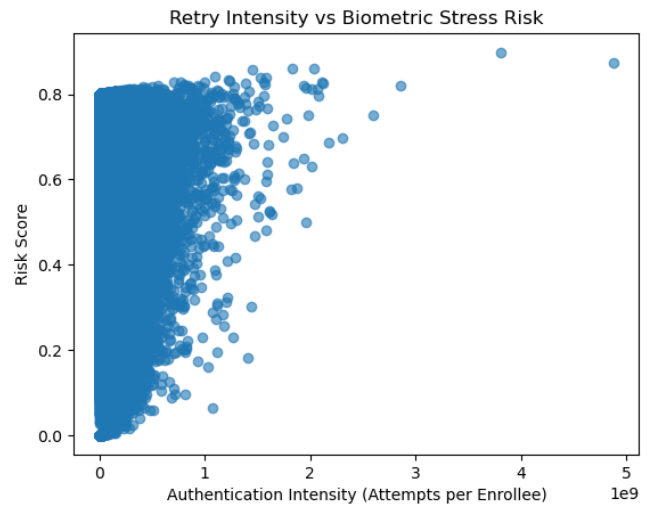


Fig. 4. Authentication intensity versus biometric failure risk. Elevated intensity indicates retry loops and authentication friction, contributing to higher observed biometric stress.

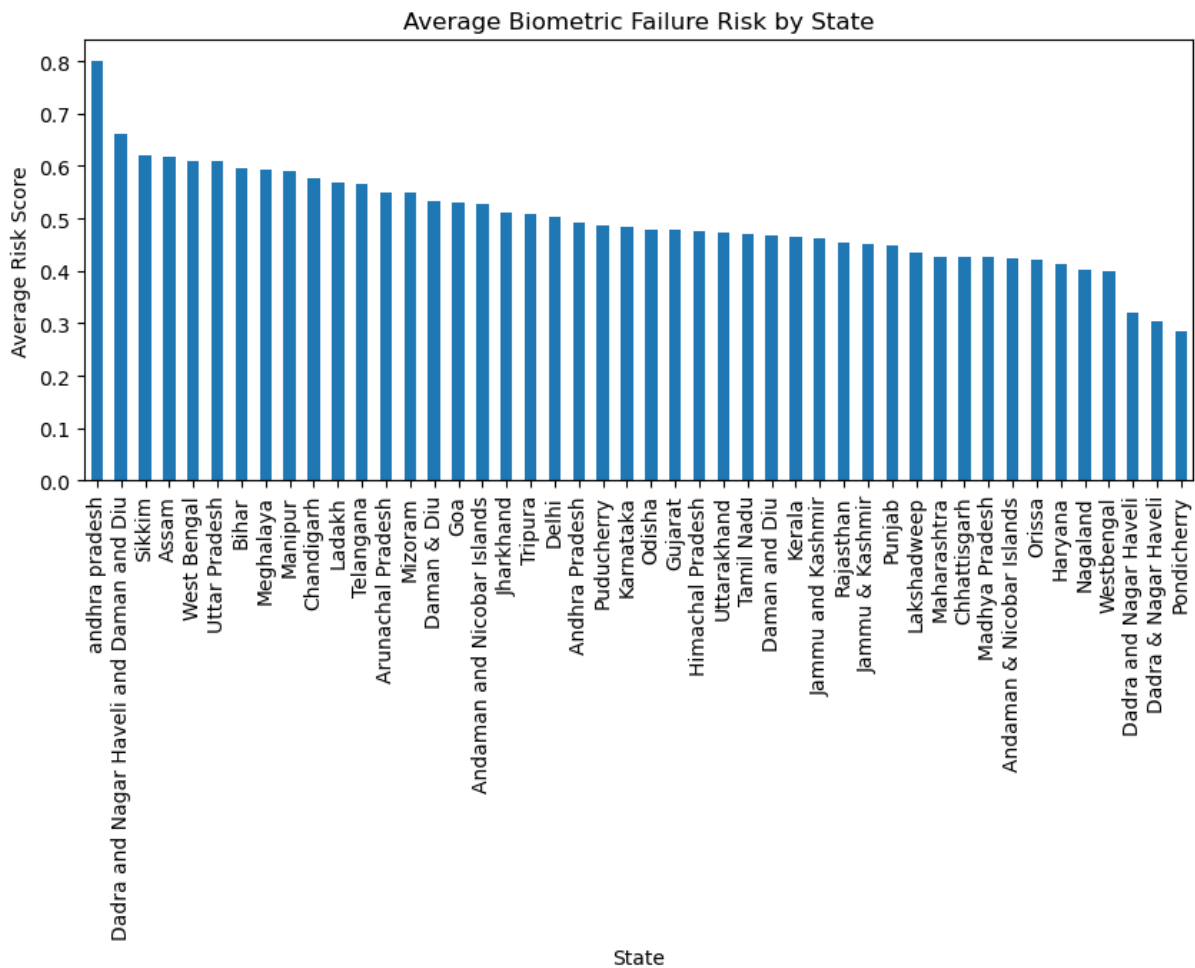


Fig. 5. Average biometric failure risk score by state. Significant regional variation is observed, enabling targeted UIDAI interventions such as infrastructure upgrades and proactive activation of alternate authentication mechanisms.