Name: Diane ISHIMWE
EMail:diane.ishimwe@amalitech.com
Specialization: Python Backend

QuickLoan Mobile Ethical Data Review

Section	Issue / Definition	Impact	Suggested Fix / Mitigation
1. Data Quality Risk	Customer data is incomplete and inconsistently formatted (e.g., missing income values, inconsistent phone formats, null employment fields before ML scoring).	Inaccurate ML loan scoring, unfair approvals/rejections, increased financial risk.	Implement mandatory field validation at mobile app level, enforce standardized formatting rules, and add automated data quality checks in preprocessing before ML scoring.
2. Legal & Compliance Risk	No explicit consent capture before collecting and storing personal data, violating lawful processing requirements under Ghana Data Protection Act (Act 843).	Legal penalties, regulatory sanctions, reputational damage, and loss of customer trust.	Implement explicit opt-in consent with privacy notice, store consent logs (timestamp + version), and conduct regular Data Protection Impact Assessments (DPIA).
Data Classification	Sensitive	High risk if breached due to financial and personal data exposure.	Apply encryption at rest and in transit, strict access control, and defined retention policies.

3. Bias & Fairness Risk	Automated ML model may produce discriminatory outcomes due to biased historical data or proxy variables (e.g., location, device type).	Unfair loan denials for certain demographic groups, ethical violations, and regulatory scrutiny.	Conduct regular bias audits, remove proxy variables, and monitor fairness metrics before and after deployment.
Source of Bias	Historical lending data and use of socioeconomic proxy variables.	Reinforces past discrimination patterns.	Apply fairness testing and retrain model with balanced datasets.
4. Storytelling / Reporting Recommendation	Implement fairness transparency reporting.	Improves accountability and builds stakeholder trust.	Publish periodic fairness reports using defined metric.
Metric to Monitor (name & definition)	Approval Rate Disparity Ratio – Ratio comparing loan approval rates between protected and non-protected groups.	Detects discriminatory approval patterns.	Review monthly and trigger investigation if disparity exceeds defined threshold.
Visualization Type (e.g., Line Chart, Grouped Bar Chart)	Grouped Bar Chart	Clearly compares approval rates across groups.	Use dashboard monitoring for leadership review.
Why It Matters (One sentence)	Ensures ethical, transparent, and non-discriminatory automated lending decisions.	Protects users and reduces regulatory risk.	Supports responsible AI governance.

DELIVERABLE 2: Corrected Data Flow Diagram (Annotated Fixes)

1. Fix Excessive Collection (Step 1 – Mobile App)

   Change: Remove access to entire contact list. Collect only necessary loan-related data (income, ID, repayment history).

Why: Applies Data Minimization principle under Ghana DPA.

2. Add Consent Layer (Between Step 2 & Step 3)

   Change: Implement Consent Management System before storing data in Raw Data DB.

   Why: Ensures compliance with Act 843 and lawful processing.

3. Add Data Classification & Retention (Step 3 – Raw Data DB)

   Change: Classify data as Sensitive. Apply encryption at rest and define retention period (e.g., 5 years).

   Why: Reduces breach risk and ensures regulatory compliance.

4. Define Handling in Preprocessing (Step 4)

   Change: Add data validation, cleaning, and bias screening before ML training.

   Why: Improves model reliability and fairness.

5. Add Transparency & Logging (Step 7 – Decision Service)

   Change: Implement explainability logs (store top 3 decision factors per loan decision).

   Why: Enables accountability and customer appeals.

6. Apply Masking & Anonymization (Step 9 & 10)

   Change: Remove direct identifiers before data enters Analytics DB or third-party systems.

   Why: Protects PII from misuse and unauthorized exposure.

DELIVERABLE 3: Summary of Review Process

As an Independent Data Governance Consultant, I applied Data Lifecycle principles to systematically evaluate QuickLoan's flawed data pipeline. I reviewed each stage of the lifecycle data collection, storage, processing, decision-making, and sharing to identify weaknesses in governance controls. By mapping risks across the lifecycle, I detected excessive data collection at the mobile app stage, absence of consent before storage, lack of classification in the raw database, and missing transparency in automated decisions.

Using Data Classification principles, I identified that the collected information qualifies as Sensitive data under Ghana's Data Protection Act (Act 843), particularly financial and personal identifiers. This classification revealed major compliance risks, especially due to missing consent mechanisms and undefined retention policies. Additionally, reviewing the preprocessing and ML stages allowed me to identify potential algorithmic bias arising from historical data and proxy variables.

To strengthen ethical governance, I proposed the Approval Rate Disparity Ratio as a core monitoring metric. This metric compares loan approval rates across demographic groups and helps detect discriminatory outcomes early. Visualizing this through a grouped bar chart allows management to identify disparities clearly and take corrective action.

By combining lifecycle analysis, data minimization, consent management, classification, and fairness monitoring, this governance framework improves compliance, transparency, and ethical accountability.