**📊 Gap Analysis Report**

**Project Name**: Customer Loan Approval & Churn Prediction  
**Prepared by**: PRATHAMESH SURVE.  
**Date**: 5-April-2025

**1. Introduction**

* **Purpose**:  
  Identify operational inefficiencies in loan approvals and customer retention processes. Recommend improvements to decision logic, data readiness, and predictive accuracy.
* **Data Sources**:
  + Loan Prediction Dataset (Train + Test)
  + Bank Customer Churn Dataset

**2. Current State Analysis**

**🏦 Loan Approval System**

* Approvals primarily rely on basic applicant information (income, credit history, etc.).
* Inconsistent approval decisions observed for applicants with similar profiles.
* Credit history is a strong influencing factor but not always available.

**💔 Customer Churn System**

* Limited visibility into behavioural patterns before churn (e.g., engagement drop-offs).
* Active members still churn without clear indicators.
* No loan-related data linked to churn risk analysis.

**3. Identified Gaps**

**⛔ Loan Process**

* **Missing Features**: No behavioural or repayment history included in prediction.
* **Data Gaps**: Incomplete or missing credit history for many records.
* **Bias Risk**: Over-reliance on credit history may lead to unjustified denials.

**🚪 Customer Retention**

* **Disconnected View**: No linkage between loan customers and churn data.
* **Predictive Gaps**: Churn reasons (e.g., dissatisfaction, poor support) not captured in data.
* **Unbalanced Data**: Churn class imbalance affecting model precision.

**4. Root Cause Analysis**

* **Siloed Data Systems**: Loan and churn datasets operate independently; lack of integrated view.
* **One-Dimensional Inputs**: Loan approval decisions made on static application data without tracking behavior or repayment patterns.
* **No Feedback Loop**: Loan outcomes (defaults, successful repayments) are not looped back for future approval optimization.

**5. Recommendations & Action Plan**

**✅ Data Integration**

* Combine churn and loan datasets for full customer lifecycle analysis.
* Add derived features like average repayment time, missed EMIs (if available), or product usage.

**✅ Model Enhancement**

* Introduce behavioural and engagement metrics in churn prediction.
* Engineer new features like tenure-to-age ratio, income-to-loan ratio.

**✅ Strategic Improvements**

* Automate risk profiling and pre-approval suggestions using model insights.
* Create early warning churn risk scores to trigger proactive engagement.

**6. Conclusion**

**🔎 Key Findings**

* Fragmented data and simple input logic limit both predictive systems.
* Major gaps in behavioural insight and customer segmentation.
* Need for unification of datasets and enhancement of features.

**🚀 Next Steps**

* Merge datasets, explore overlaps, and define common customer profiles.
* Build a pipeline that captures loan approval outcomes and churn events in one system.
* Develop dashboards for real-time monitoring of loan approval efficiency and churn risk zones.