# . Problem Overview

FinFast is a digital lending fintech offering short-tenure loans across Retail, MSME, and Consumer Durable segments. The company aims to enhance underwriting efficiency and better understand borrower behaviour through advanced predictive modelling. The objective is to increase repeat-loan conversion and strengthen overall credit decisioning.

# 2. Candidate Task

You are required to:

- Build baseline and advanced machine learning models to predict:
  - Repeat-loan probability
- Identify key drivers, quantify business impact, and provide clear recommendations for leadership through a concise business summary.

# 3. Dataset Description

A synthetic dataset with **100,000 loan applicant records** is provided. It contains demographic details, credit bureau information, behavioural transaction metrics, and loan performance outcomes.

The dataset may include real-world challenges such as missing values, inconsistent entries, noise, skewed distributions, imbalanced labels, and outliers.

| Column Name | Description |
| --- | --- |
| customer_id | Unique customer identifier |
| age | Customer age in years |
| gender | Customer gender (Male / Female / Other / Blank) |
| city | Customer city (may contain inconsistent spellings) |

| Column Name | Description |
|---|---|
| **income** | Reported monthly income (mixed units possible) |
| **employment_type** | Employment category (Salaried / Self-Employed / Other) |
| **credit_score** | Credit bureau score (simulated; 300–900) |
| **loan_amount** | Applied loan amount (INR) |
| **approved_amount** | Amount approved (INR; may include manual adjustments) |
| **application_date** | Date of loan application |
| **registration_timestamp** | Account registration timestamp |
| **nb_txn_cnt** | Number of net-banking transactions in a recent window |
| **days_since_last_txn_credit_repayment_24month** | Days since last credit repayment transaction in last 24 months |
| **application_channel** | Acquisition channel (Organic / Paid / Referral) |
| **loan_purpose** | Purpose of loan (Medical / Education / Travel / Other) |
| **device_type** | Device used for application (Android / iOS / Other / Blank) |
| **text_complaint** | Short free-text complaint (may be empty/null) |
| **repeat_loan** | Target: 1 = took another loan within 12 months, 0 = otherwise |

# 5. Business Decision-Making Expectations

1. Identify key drivers affecting repeat-loan behaviour.
2. Recommend credit and customer-segmentation strategies.
3. Suggest retention actions for high-value customers.
4. Highlight major risks (data drift, bias, leakage) if any.
5. Quantify business impact of model performance shifts (e.g., recall drop).