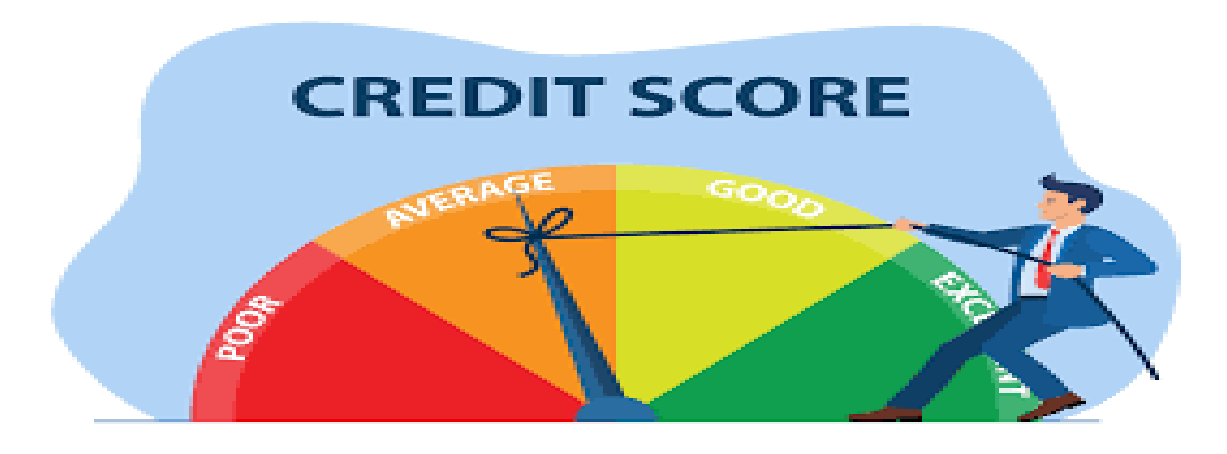
Smart Credit Risk Scoring Using Machine Learning

Predicting Defaults to Empower Ethical Lending

A Non-Technical Presentation



Overview

- Rising loan default rates threaten the sustainability of a digital micro-lender in Kenya.
- Current manual rule-based credit scoring lacks precision.
- We used machine learning to build a data-driven credit risk model.
- This project supports smarter, fairer lending and financial inclusion.

Business and Data Understanding

- Business Goal: Predict if a client will default next month.
- Data: Credit history, behavior, and demographics of 30,000+ clients.
- Stakeholders: Executives, risk officers, and product developers.
- Benefit: Early identification of high-risk clients to manage risk and increase efficiency.

Modeling

- We trained classification models to predict loan default:
  - Baseline Logistic Regression
  - Logistic Regression with SMOTE
  - Random Forest with SMOTE
- Addressed class imbalance using SMOTE (Synthetic Minority Over-sampling Technique).
- Random Forest model performed best overall.

Evaluation

- Key Metrics for Default Class (1):

  - Accuracy: 91% → Improved to 97% with threshold tuning.

  - F1-Score: Increased from 0.34 to 0.52 with optimal threshold.

  - AUC Score: 0.91 → Strong model discrimination.

- Top Features: Bank Balance, Index(Credit score), Credit history, Annual Salary and Employment status.

Recommendations

- Deploy Random Forest model to screen clients.
- Integrate the model into loan approval workflow.
- Use top features to develop early intervention strategies.
- Monitor model performance and retrain regularly.
- Integrating additional behavioral and transactional data (e.g mobile money usage, repayment history).

Next Steps

- Pilot the model on a subset of applications.
- Engage with product and compliance teams for integration.
- Collect feedback and iterate.
- Scale deployment across operations.
- Train model with new customer data for adaptability

Thank You

- We welcome your questions and look forward to discussing how this model can support ethical, data-driven lending at scale.

By: Allan Ofula

LindedIn: https://www.linkedin.com/in/allanofula

Portfolio: https://allan-ofula.github.io/#projects

Github: https://github.com/Allan-Ofula

Email: ofulaallan@gmail.com