LENDSMART CREDIT RISK ANALYSIS

LUIS ALAN MORALES CASTILLO A01659147
PAULINA DÍAZ ARROYO A010295932
RODRIGO JIMÉNEZ ORTIZ A01029623

Business Problem

LendSmart currently faces a 28% loan default rate, which management considers unsustainably high. Every defaulted loan represents a direct financial loss, and the company needs a reliable way to identify high-risk applicants before approving a loan.

However, there is a critical trade-off:

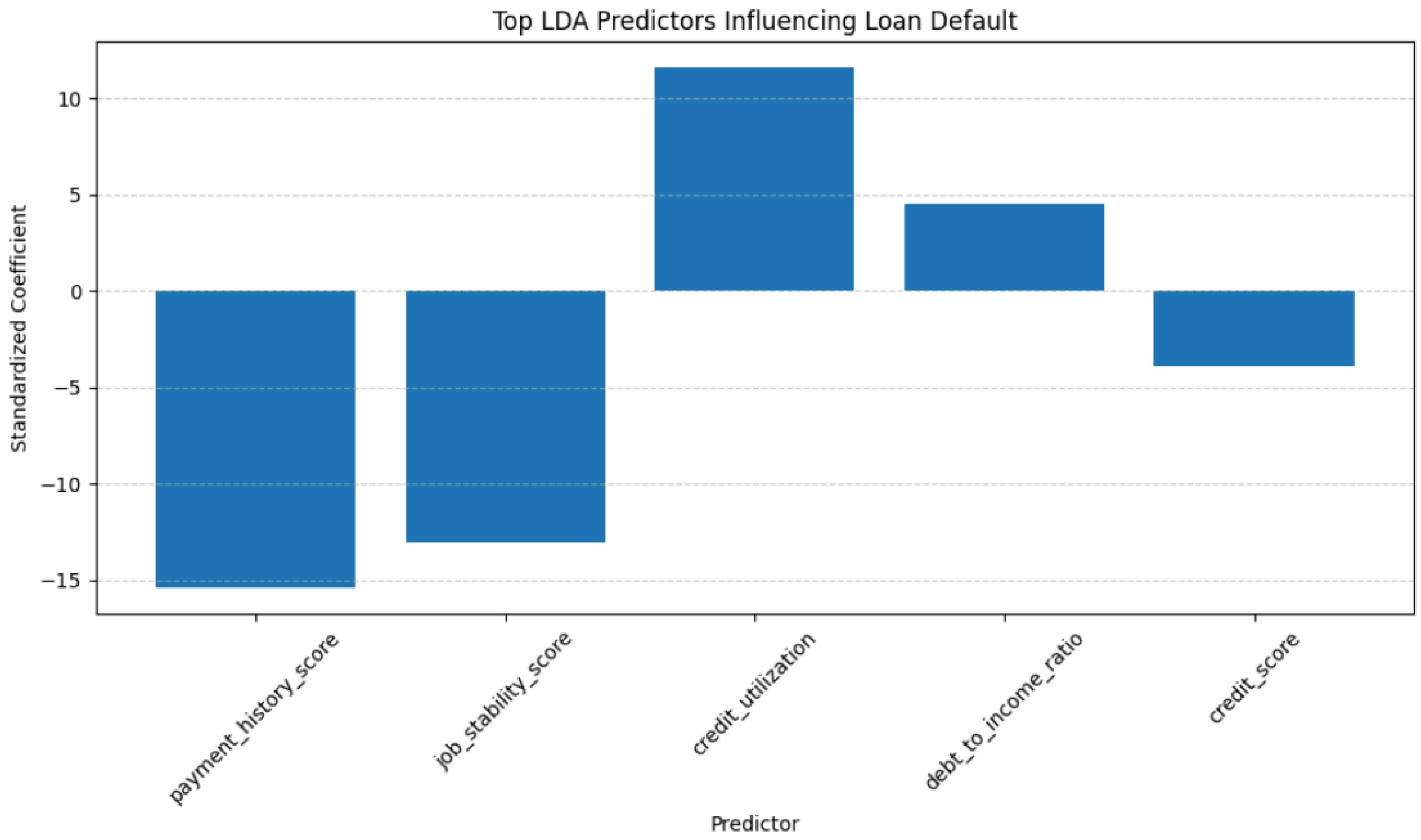
- Approving a “bad” loan leads to immediate financial loss. This is the worst outcome.
- Rejecting a “good” loan results in lost revenue and a dissatisfied potential customer.

Our objective is to build a predictive model that strikes the right balance. Minimize missed defaulters while keeping unnecessary rejections under control.

Key Insights

Our analysis shows the five strongest financial drivers of loan default risk, ranked from most to least influential:

- **Payment History:** Borrowers with missed or late payments on previous credit obligations are far more likely to default.
- **Job Stability:** Applicants with frequent job changes or short employment histories face higher default risk due to uncertain income and limited ability to manage financial shocks.
- **Credit Utilization:** High utilization of existing credit lines indicates borrowers are heavily reliant on credit, leaving little room to absorb new loan payments.
- **Debt-to-Income Ratio:** A high ratio means a large portion of income is already committed to debt, reducing the borrower’s capacity to take on additional obligations.
- **Credit Score:** Lower scores remain a strong indicator of default risk, reflecting past patterns of late payments, credit mismanagement, and financial strain.



Model Performance: LDA vs QDA

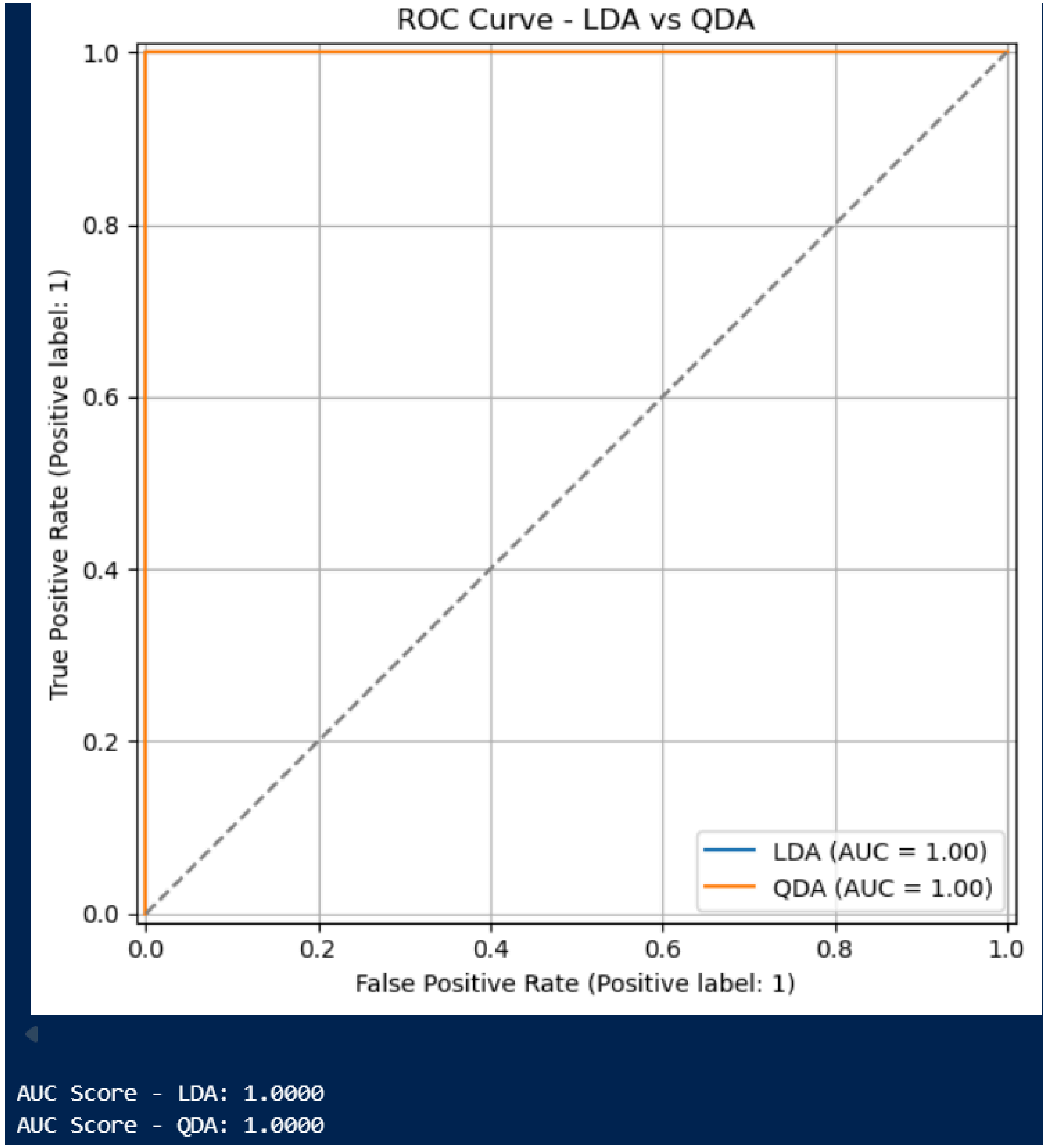
We compared two statistical models designed to predict loan default:

- LDA (Linear Discriminant Analysis)
- QDA (Quadratic Discriminant Analysis)

The ROC curve shows their predictive accuracy on unseen data.

- The closer the curve is to the top-left corner, the better the model separates defaulters from non-defaulters.
- The AUC score summarizes this performance.

In our results, both models achieved perfect classification, with identical AUC values of 1.0. This means that every default and non-default case was correctly identified. Although both models performed equally well with this dataset, LDA may be preferred because it is simpler, more stable, and easier to interpret.



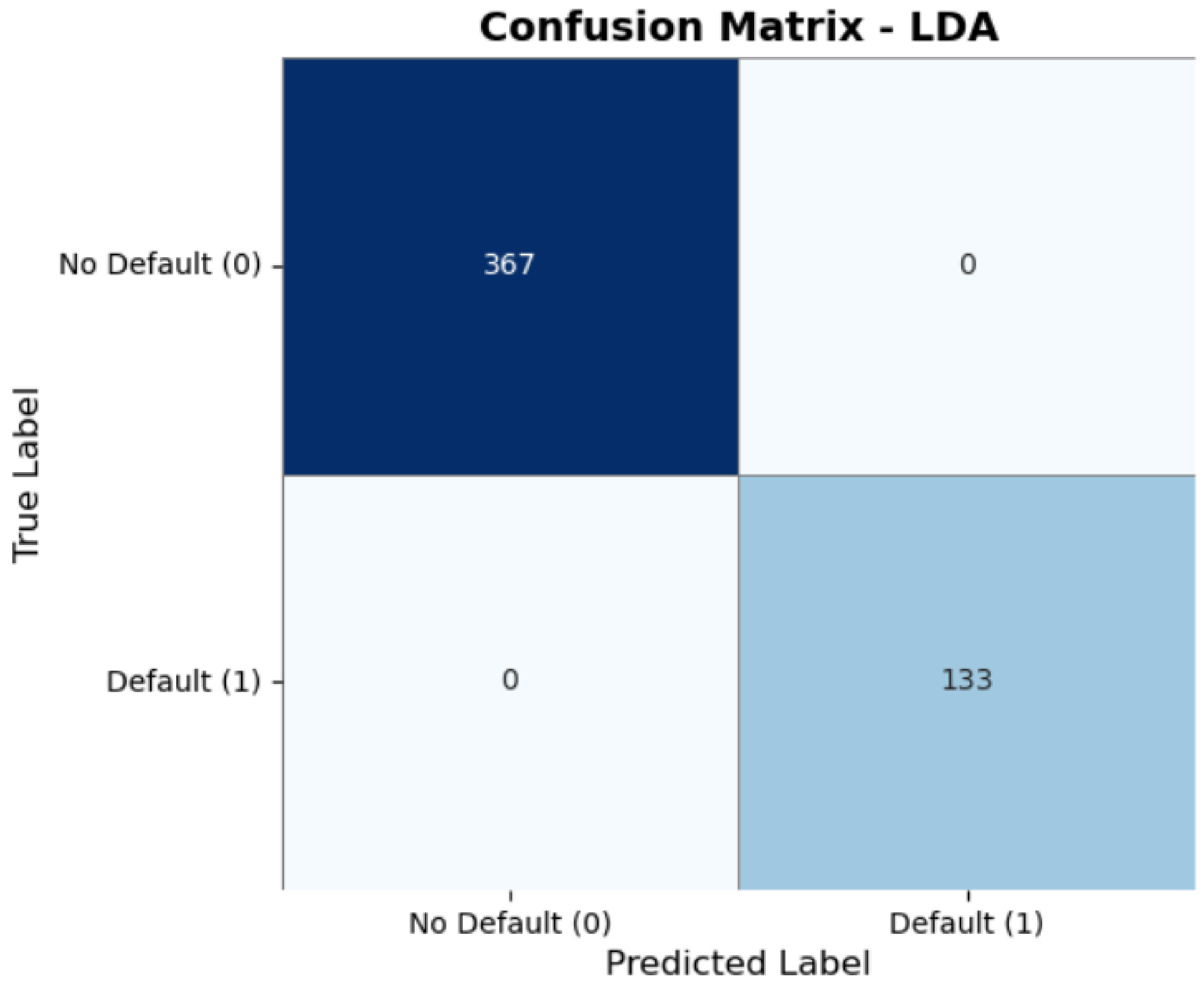
Business Trade-Off

We can interpret the model’s results as follows:

- 367 good customers correctly approved (true negatives)
- 133 high-risk customers correctly flagged (true positives)
- 0 good customers wrongly rejected (false positives)
- 0 defaulters missed (false negatives)

Business Impact:

- Correctly identifies 100% of all actual defaulters.
- Maintains 100% overall accuracy.
- Keeps false rejections filtered.



Final Recommendations

Recommended Next Step:
Integrate the model into LendSmart’s loan evaluation workflow and begin a pilot program to monitor real-world performance.

Thank you

VIDEO LINK:
PRESENTATION VIDEO