<u>Fraud Detection</u>

Problem Framing:

<table>
<tr><th></th><th>Qualitative</th><th>Quantitative</th><th>Question</th></tr>
<tr><td>Current State</td><td>Increase in fraudulent transaction => bad customer experience => less customers => less revenue => loss to the bank</td><td>10% fraudulent transactions => 5% less customers => 5% less revenue</td><td>What is the average number of fraudulent transactions at present and what can be done about it in order to decrease them?</td></tr>
<tr><td>Objectives</td><td>Build a model that can detect fraudulent transactions
Less fraud => better customer experience => increase revenue</td><td>Find and decrease the fraudulent transactions.</td><td>How do we detect fraudulent transactions?</td></tr>
<tr><td>Benefit/Cost Tradeoff and Prioritization</td><td>Errors -

TP - Fraud identified => customers protected => better user experience => more revenue to bank

FP – Valid transaction marked fraud => bad user experience => less revenue

FN - Fraud marked valid => Risk to customers’ assets => bad user experience => Less revenue

TN - Valid transaction marked valid => no significant impact on revenue</td><td>cost-benefit matrix
<table><tr><td>c(TP)</td><td>c(FP)</td></tr><tr><td>c(FN)</td><td>c(TN)</td></tr></table>
Improve the predictive model with one that yields a lower number of false positives AND a lower number of false negatives. In other words, improve the precision AND the recall.</td><td>What are the costs of errors/benefits of correct predictions and why?</td></tr>
<tr><td>Constraints</td><td>Can only afford very little FN rate and less FP rate</td><td>At most 5% FN and 10% FP=> Customer risk and better user experience</td><td>What are the acceptable risks/budgets and why?</td></tr>
</table>

Desired State	• Benefit: significantly lesser fraudulent transactions => significantly better user experience => significantly more customers => significantly better revenue • Cost: very few false negatives => limited risk of bad user experience => limited risk of losing customers => limited risk to revenue	• at least 50% decrease in fraudulent transactions (from 20% to 10%) => 5% better engagement => 5% more revenue • at most 10% false positives => 1% loss of customer =>0.1% risk to revenue	What is the desired outcome (benefits/costs) that we want to see and why?

Why ML

	qualitative	quantitative	question
best non-ML alternative hypothesis	classify based on amount or location of transaction => too many FP and FN => more fraud and bad user experience => lesser customers => loss of revenue	50% FP 70% FN => not cleaning enough fraudulent transactions and causing more complaints for misclassifying genuine transactions as fraudulent => 5% revenue loss risk	What are the non-ML alternatives and why are they problematic? (pains/missed gains)?
ML value proposition hypothesis	much fewer FP and FN => better user experience => more revenue	10% FP 50% FN => 50% less fraud in expense of 1% bad engagements => 5% increase in revenue at the expense of 0.1% risk	What are the advantages (pain relievers/gain creators) of ML solutions and why?
ML feasibility hypothesis	• data: labeled dataset of each person's bank history • model: state of the art review suggests	• data: around five thousand samples • model: state of the art claim solutions with 10% FP 20%	What data and models are good candidates and why?

	promising candidates are available	FN	

ML Solution Design

	choices	metrics	experiment
data	(labeled) transaction data	• label imbalance	• randomized 70/15/15 train/validation /test split
model	pr(fraud)	• AUCPR (Precision recall curve)	• rule based heuristic • tf-idf + logistic regression • tf-idf + random forest • BERT + logistic regression train these benchmark models using train data. validate and tune using validation data. select the model with best AUCPR on test data
action	if Pr(fraud) > threshold: auto take down	• precision • recall • confusion matrix	• choose a threshold to maximize the recall (estimated reward) subject to precision > 90%
reward	• decrease in fraud • cost of misclassification	• % Decrease in fraud • % Increase in daily active users	• A/B test