Lecture 5: 22 January, 2026

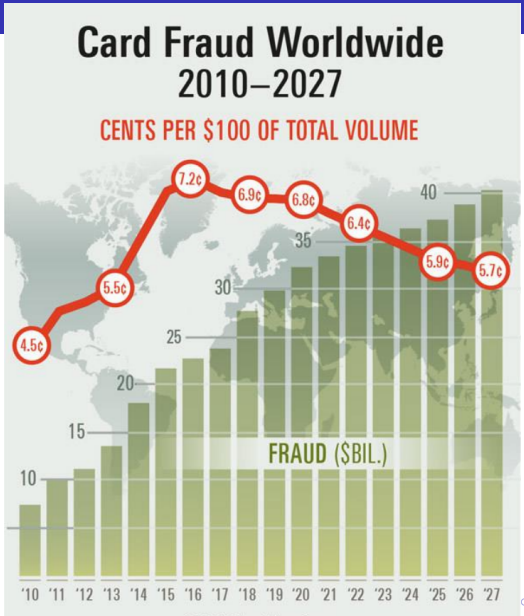
Madhavan Mukund

https://www.cmi.ac.in/~madhavan

Data Mining and Machine Learning
January–April 2026

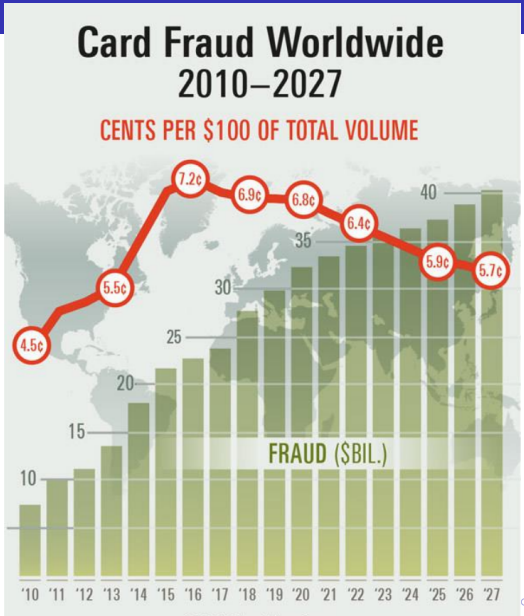
What are we measuring?

- Accuracy is an obvious measure
  - Fraction of inputs where classification is correct
- Classifiers are often used in asymmetric situations
  - Less than 1% of credit card transactions are fraud
- "Is this transaction a fraud?"
  - Trivial classifier — always answer "No"
  - More than 99% accurate, but useless!



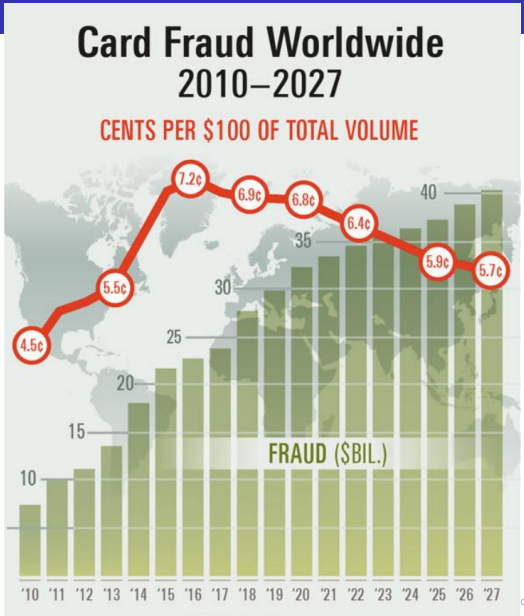
Catching the minority case

- The minority case is the useful case
  - Assume question is phrased so that minority answer is "Yes"
  - Want to flag as many "Yes" cases as possible



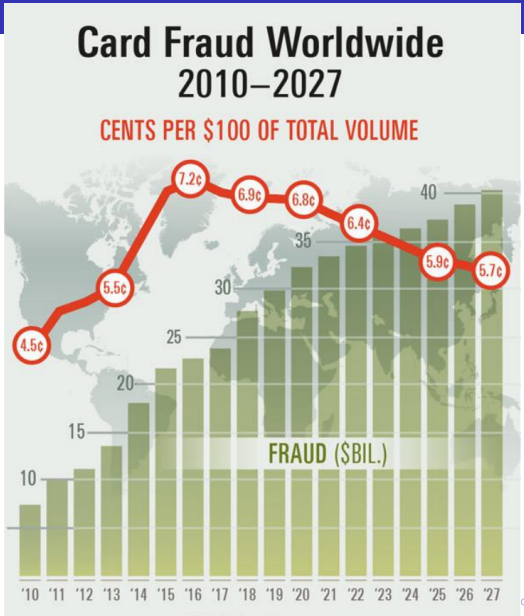
Catching the minority case

- The minority case is the useful case
  - Assume question is phrased so that minority answer is "Yes"
  - Want to flag as many "Yes" cases as possible
- Aggressive classifier
  - Marks borderline "No" as "Yes"
  - False positives



Catching the minority case

- The minority case is the useful case
  - Assume question is phrased so that minority answer is "Yes"
  - Want to flag as many "Yes" cases as possible
- Aggressive classifier
  - Marks borderline "No" as "Yes"
  - False positives
- Cautious classifier
  - Marks borderline "Yes" as "No"
  - False negatives



Confusion matrix

- Four possible combinations
  - Actual answer: Yes / No
  - Prediction: Yes / No

Confusion matrix

- Four possible combinations
  - Actual answer: Yes / No
  - Prediction: Yes / No
- Record all four possibilities in confusion matrix
  - Correct answers
    - True positives, true negatives
  - Wrong answers
    - False positives, false negatives

	Classified positive	Classified negative
Actual positive	True Positive (TP) ✓	False Negative (FN) ?
Actual negative	False Positive (FP) ?	True Negative (TN) ✓

Performance measures

Precision - Recall Tradeoff

Precision

- What percentage of positive predictions are correct?

$$\frac{\text{TP}}{\text{TP} + \text{FP}}$$

Recall

- What percentage of actual positive cases are discovered?

$$\frac{\text{TP}}{\text{TP} + \text{FN}}$$

	Classified positive	Classified negative
Actual positive	True Positive (TP) ✓	False Negative (FN)
Actual negative	False Positive (FP) ✗	True Negative (TN)

Performance measures

- Precision 1, Recall 0.01

Yes = $\frac{100}{1000}$ = 10%

	Classified positive	Classified negative
Actual positive	1	99
Actual negative	0	900

Performance measures

- Precision 1, Recall 0.01
- Recall up to 0.4, but precision down to 0.29

	Classified positive	Classified negative
Actual positive	40	60
Actual negative	100	800

Performance measures

- Precision 1, Recall 0.01
- Recall up to 0.4, but precision down to 0.29
- Recall up to 0.99, but precision down to 0.165

	Classified positive	Classified negative
Actual positive	99	1
Actual negative	500	400

Performance measures

- Precision 1, Recall 0.01
- Recall up to 0.4, but precision down to 0.29
- Recall up to 0.99, but precision down to 0.165
- Precision-recall tradeoff
  - Strict classifiers : fewer false positives (high precision), miss more actual positives (low recall)
  - Permissive classifiers : catch more actual positives (high recall) but more false positives (low precision)

	Classified positive	Classified negative
Actual positive	99	1
Actual negative	500	400

Performance measures

- Which measure is more useful?
  - Depends on situation
- Hiring
  - Screening test: high recall
  - Interview: high precision
- Medical diagnosis
  - Immunization: high recall
  - Critical illness diagnosis: high precision

	Classified positive	Classified negative
Actual positive	True Positive (TP)	False Negative (FN)
Actual negative	False Positive (FP)	True Negative (TN)

Performance measures

Other measures, terminology

- Recall is also called sensitivity
- Accuracy: (TP+TN)/(TP+TN+FP+FN)
- Specificity: TN/(TN+FP)
- Threat score: TP/(TP+FP+FN)
  - TN usually majority, ignore, not useful

	Classified positive	Classified negative
Actual positive	True Positive (TP)	False Negative (FN)
Actual negative	False Positive (FP)	True Negative (TN)

Performance measures

Other measures, terminology

- Recall is also called sensitivity
- Accuracy: (TP+TN)/(TP+TN+FP+FN)
- Specificity: TN/(TN+FP)
- Threat score: TP/(TP+FP+FN)
  - TN usually majority, ignore, not useful

	Classified positive	Classified negative
Actual positive	True Positive (TP)	False Negative (FN)
Actual negative	False Positive (FP)	True Negative (TN)

F Score

- A single combined score
- Harmonic mean of precision, recall

$$\frac{2pr}{p+r}$$