① Mindbridge Overview

② Fraud as Class Imbalance

③ Fraud Algos

④ Real World Learnings — what makes

Mind Bridge.

~$200B financial loss/annum
→ what we know about.

~$4 Trillion/annum
→ estimated loss of "unknown".
error/intent.

State of Fraud
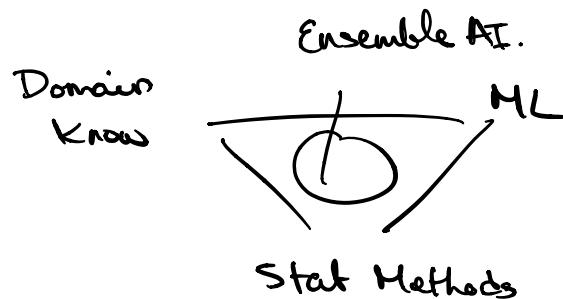
40% — tip line or by mistake
03% — by analytics — due to sampling
and only using [1,3]% ledger.
16 months to detect fraud
80% surge in fraud 2016
• 0.1% of trx are fraud
→ class imbalance
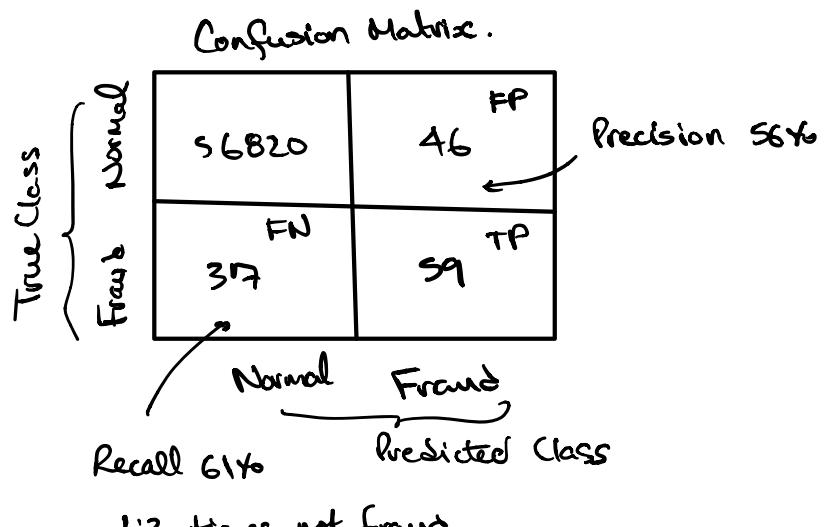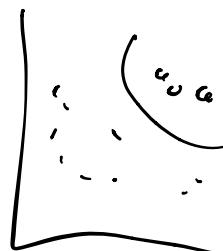
• People make more errors than frequency of fraud.

Ensemble AI.
Domain Know   ML
Stat Methods
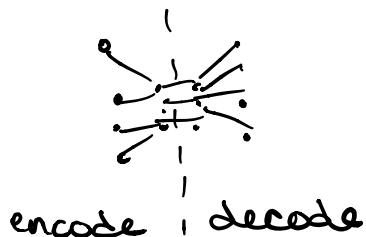
Algorithm Selection

Logistic Regression 97% Accuracy

Confusion Matrix.

		Normal	Fraud
True Class	Normal	56820	46 (FP)
	Fraud	317 (FN)	59 (TP)

Precision 56%

Normal   Fraud
Predicted Class

Recall 61%

1:2 times not fraud

(*) Class imbalance hides inaccuracy.

Autoencoder

54% Recall
14% Precision

encode | decode

encodes & reconstructs

if output doesn't match input
↳ outliers

Confusion Matrix.

		Normal	Fraud
True Class	Normal	56477	370 (FP)
	Fraud	53 (FN)	62 (TP)

Precision 14%

Normal   Fraud
Predicted Class
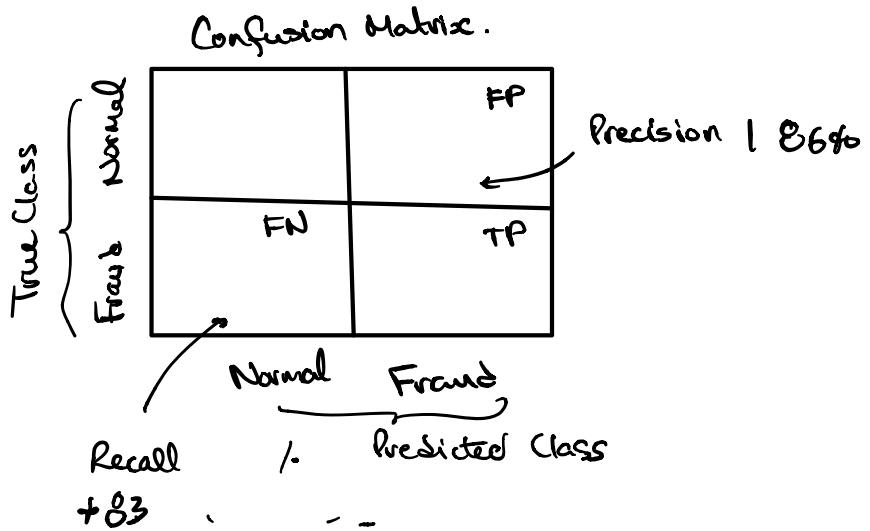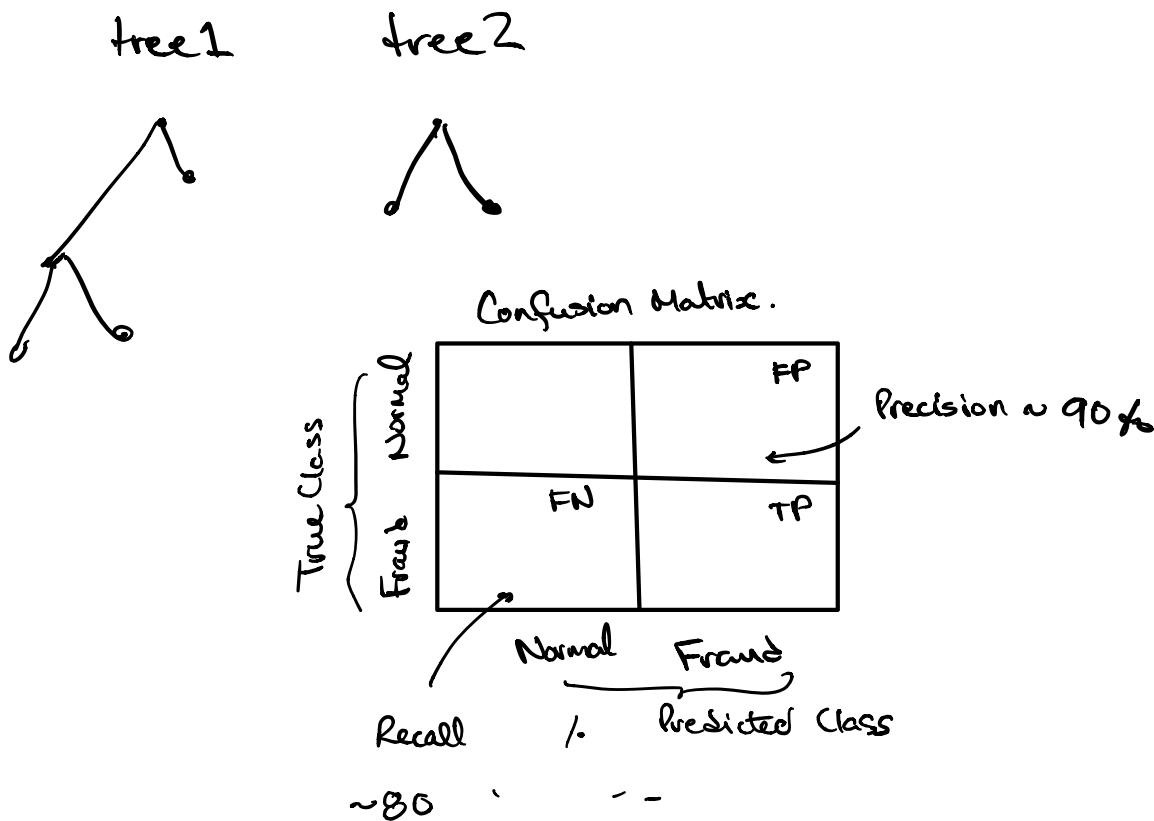
Recall 54%

1:2 times not fraud

* Stochastic outlier detection & autoencoder
  for unsupervised learning / unlabelled data.

DL   6 layers, Adam Opt. ReLU

Confusion Matrix.

		Normal	Fraud
True Class	Normal		FP
	Fraud	FN	TP

Precision | 86%

Normal   Fraud
Predicted Class

Recall
+83

Gradient Boosting

tree1   tree2

Confusion Matrix.

		Normal	Fraud
True Class	Normal		FP
	Fraud	FN	TP

Precision ~ 90%

Normal   Fraud
Predicted Class

Recall
~80

* Based on manufactured Kaggle Data.

* Could handle bad fraud data/unreliable.
            [3,10]% recall improvement

* syn. labelling → poorer results.

Feature Importance

- find good features
- join data.