① Mind Bridge 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
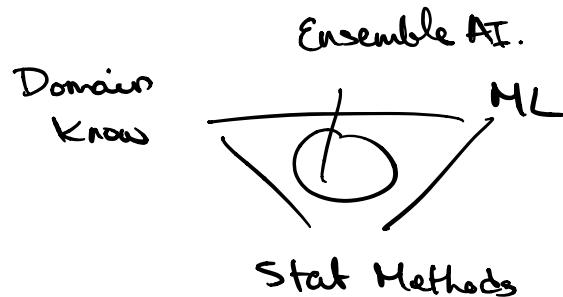
40% — tip line or by mistake
0.3% — by analytics — due to sampling
and only using [1,3]% ledger.
16 months to detect fraud
80% surge in fraud 2016
• 0.1% of trx are fraud
↳ class imbalance
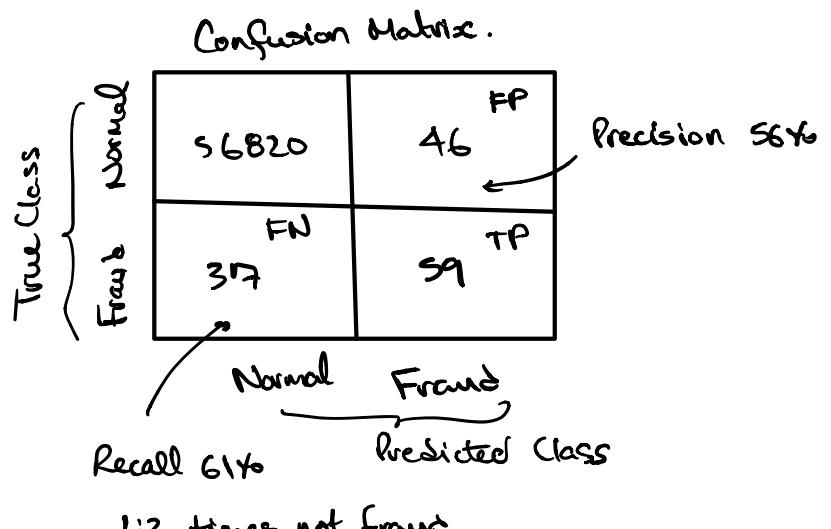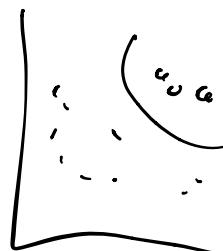
• People make more errors than frequency of fraud.

Ensemble AI.
Domain Know   ML
Stat Methods
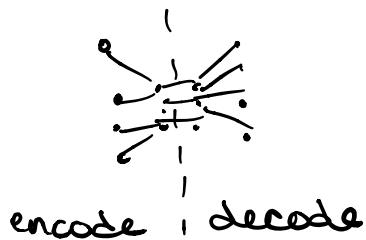
Logistic Regression 97% Accuracy

Confusion Matrix.

True Class

	Normal	Fraud
Normal	56820	46 FP
Fraud	37 FN	59 TP

Precision 56%

Normal   Fraud

Predicted Class

Recall 61%

1:2 times not fraud

(*) Class imbalance hides inaccuracy.

Autoencoder

encode | decode

54% Recall

14% Precision

encodes & reconstructs

if output doesn't match input
↳ outliers

Confusion Matrix.

True Class

	Normal	Fraud
Normal	56477	370 FP
Fraud	53 FN	62 TP

Precision 14%

Normal   Fraud

Predicted Class
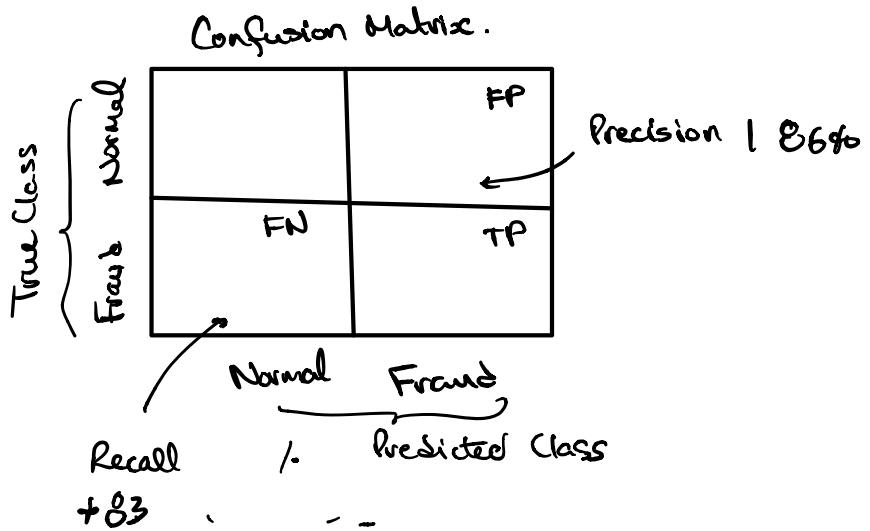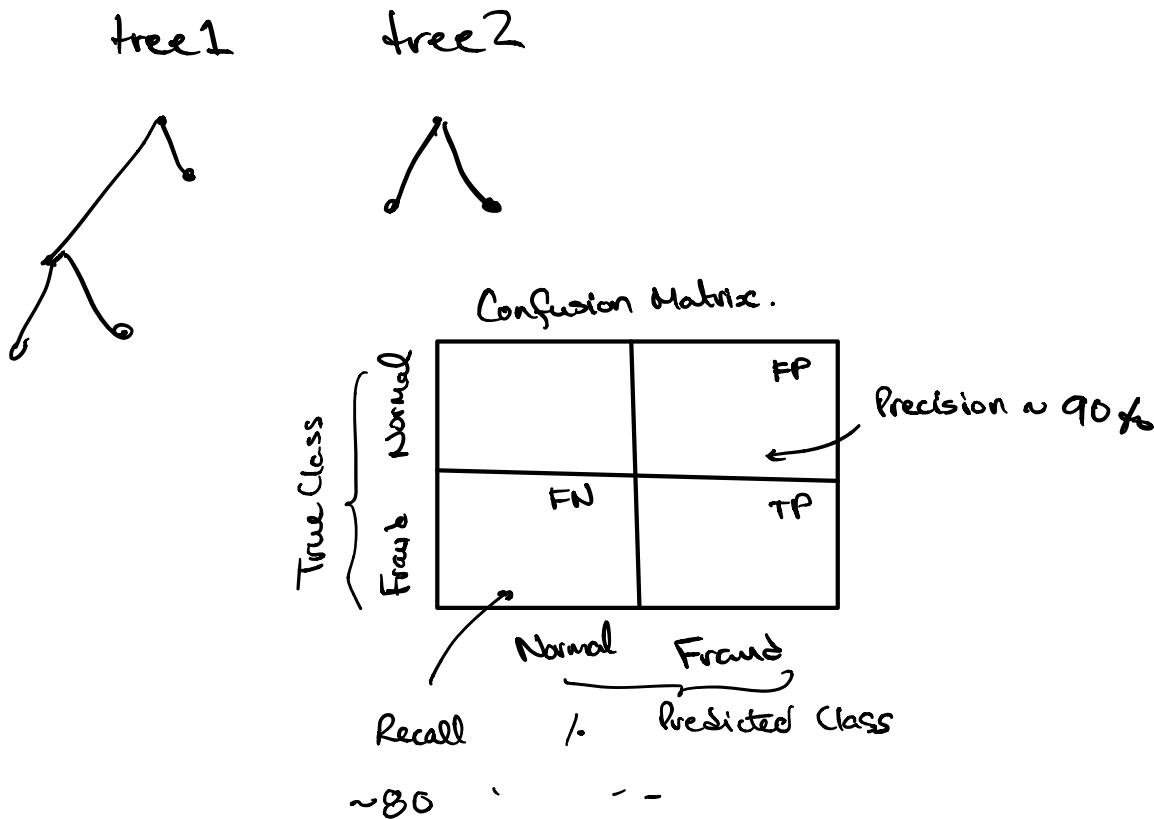
Recall 54%

1:2 times not fraud

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

DL   6 layers, Adam Opt. ReLu

Confusion Matrix.

		FP
	FN	TP

True Class (Normal / Fraud)

Normal   Fraud
Predicted Class

Precision | 86%

Recall
+83 %

Gradient Boosting

tree1   tree2

Confusion Matrix.

		FP
	FN	TP

True Class (Normal / Fraud)

Normal   Fraud
Predicted Class

Precision ~ 90%

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.