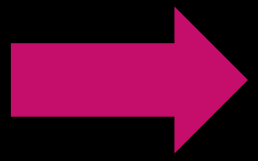
BEHAVIORAL FRAUD MODELING:

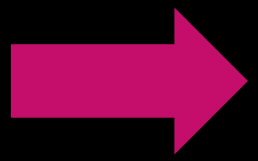
Separating Honest Patterns from Fraudulent Chaos

First-party Behavior Data



Consumer Insights

First-party Behavior Data



Consumer Insights

We want cleaner data.

Manual Check and Hard Rules

Manual Check and Hard Rules

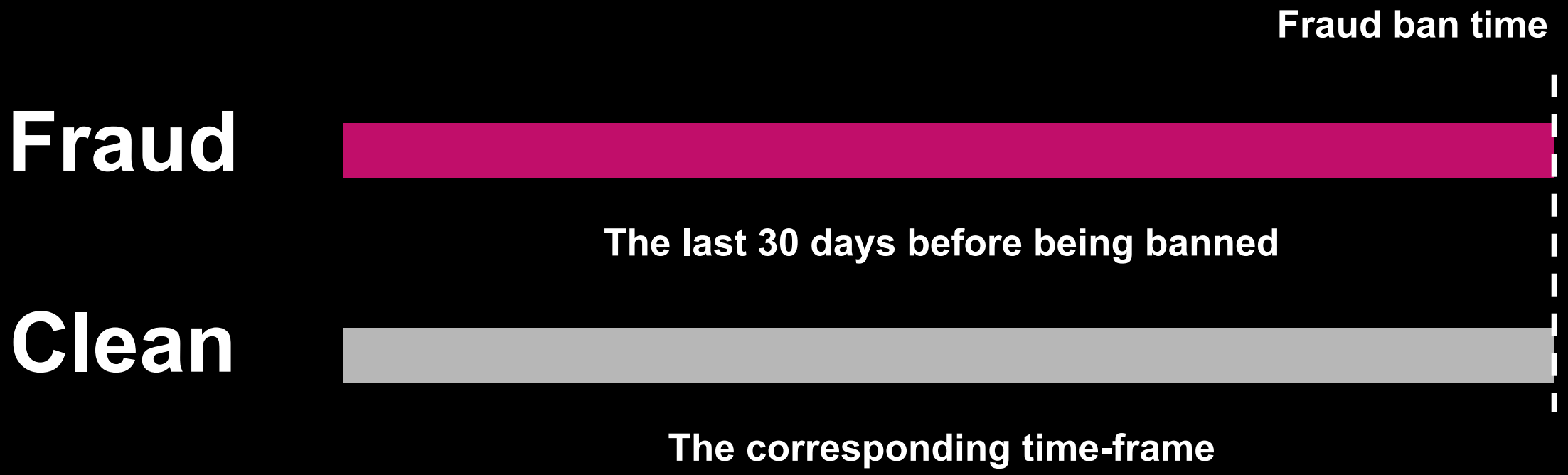
A scalable monthly screening model that outputs confidence scores

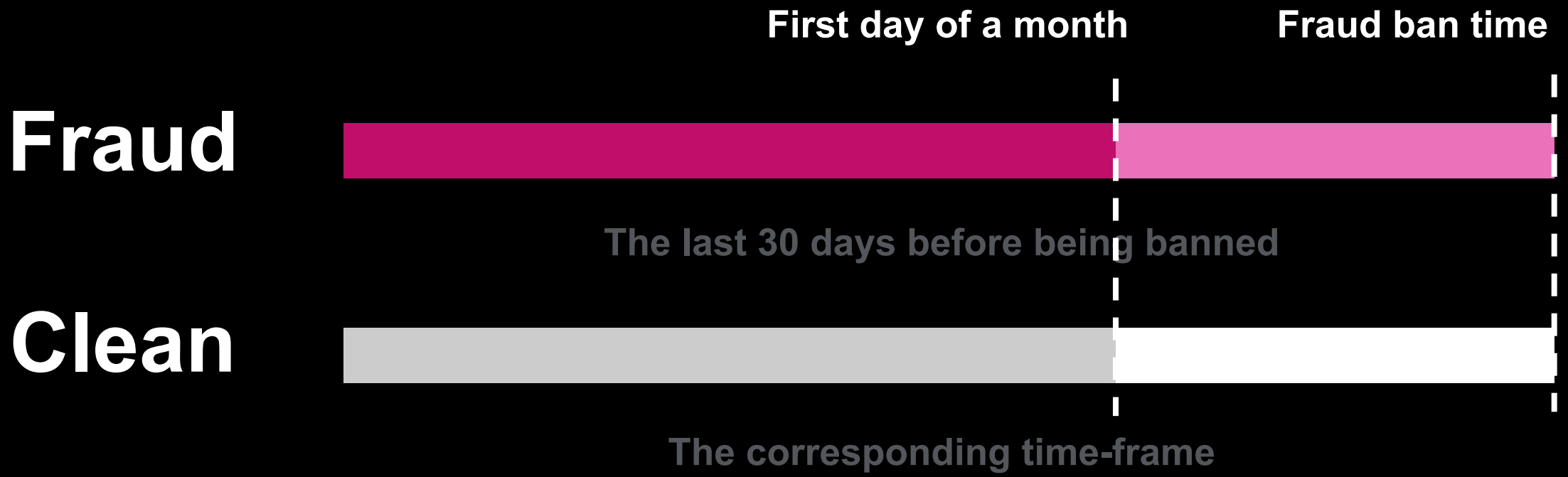
Fraud Behaves Differently

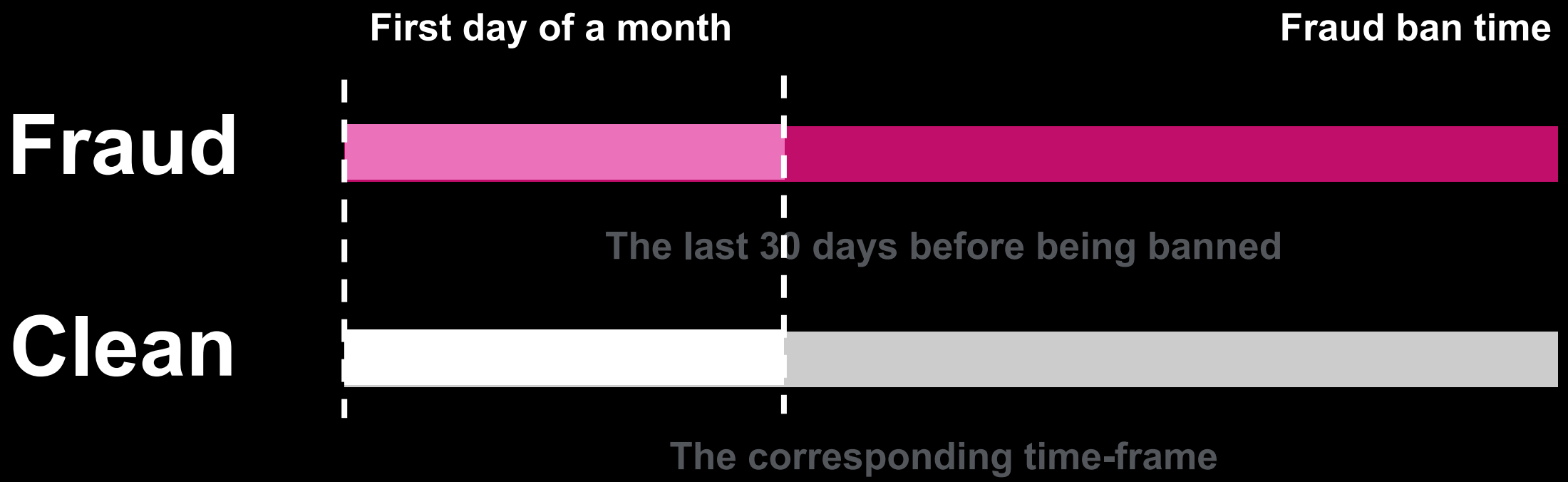
Fraud Behaves Differently

Not identity-based but pattern-based

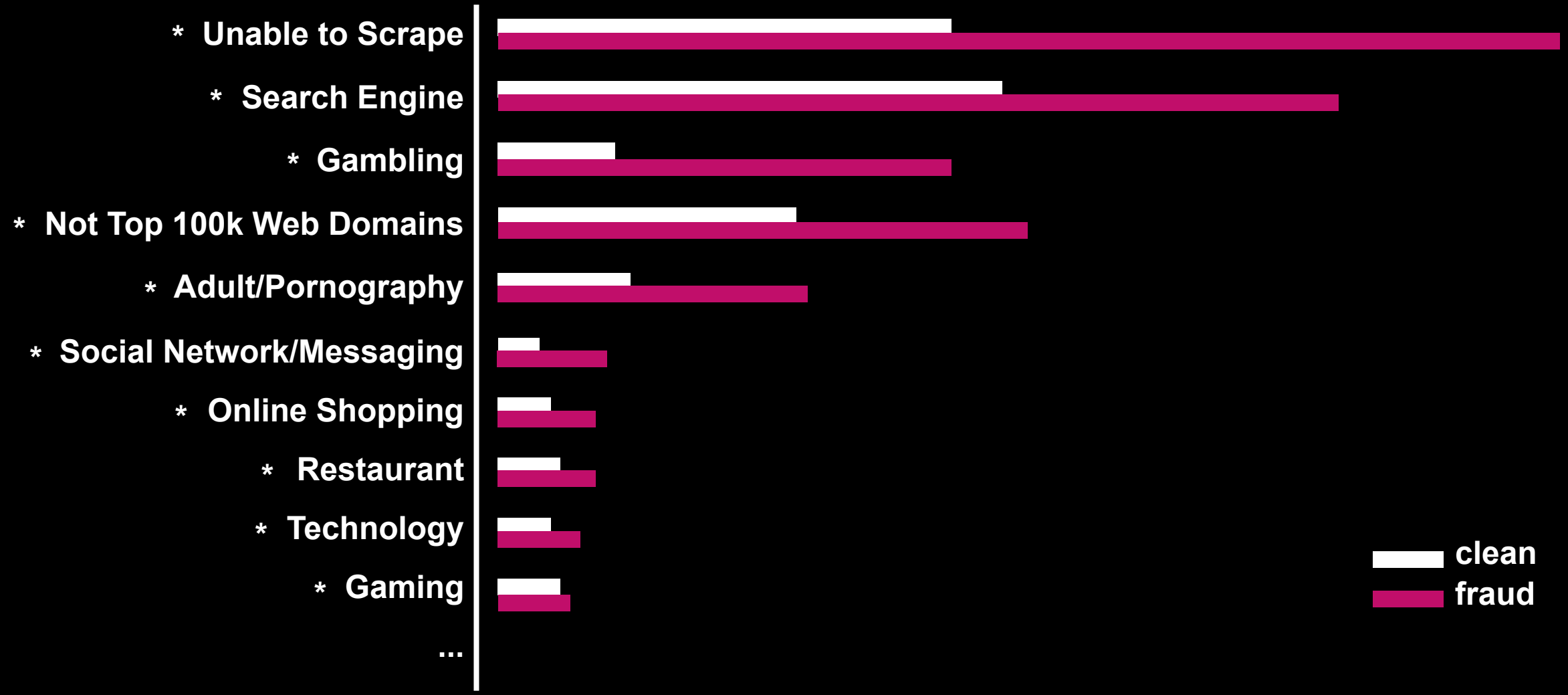
Detectable in digital traces


Fraud ban time
Fraud
The last 30 days before being banned
Clean
The corresponding time-frame

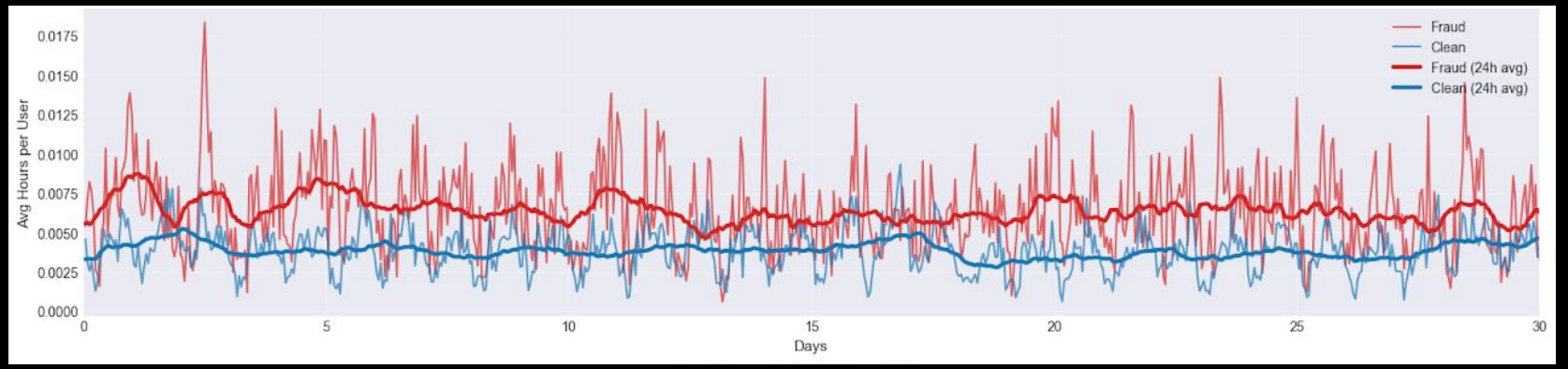

First day of a month
Fraud ban time
Fraud
The last 30 days before being banned
Clean
The corresponding time-frame


First day of a month
Fraud ban time
Fraud
Clean
The last 30 days before being banned
The corresponding time-frame

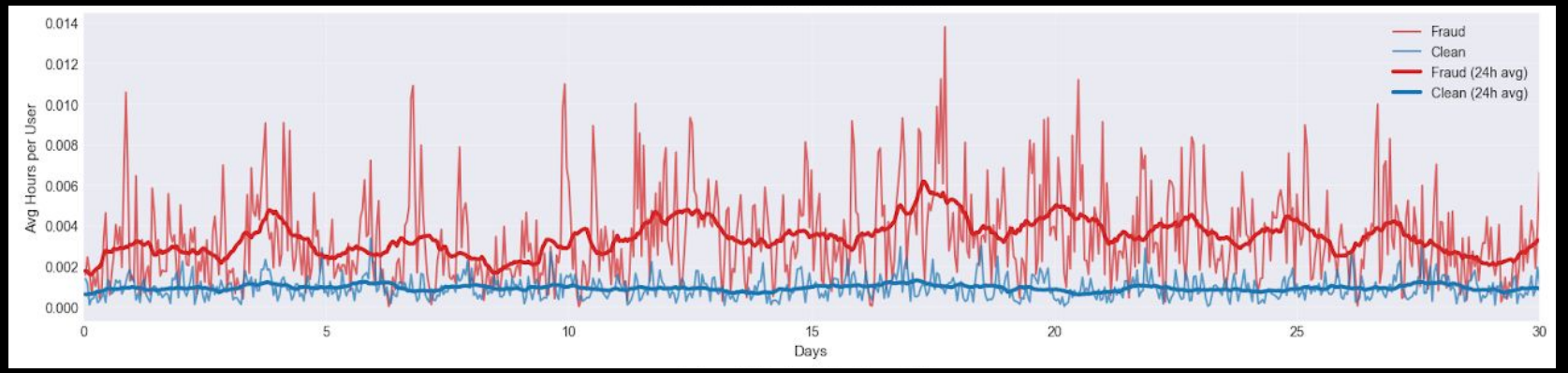
Fraud vs Clean Users – Average Time per Web Category



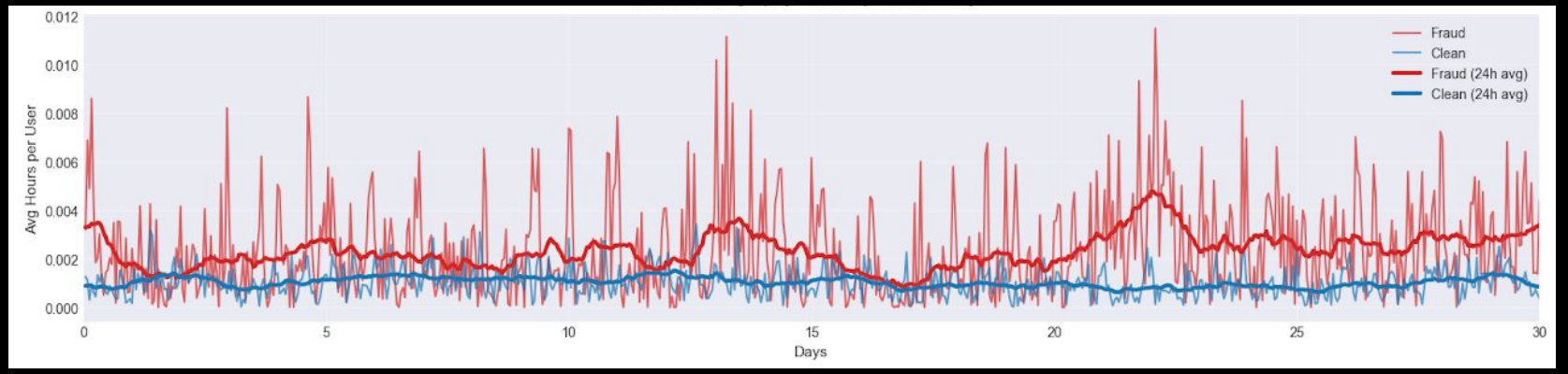
Search Engine



Gambling



Adult/Pornography



Web Activities
App Activities
Search Terms
Aggregated & Demographic

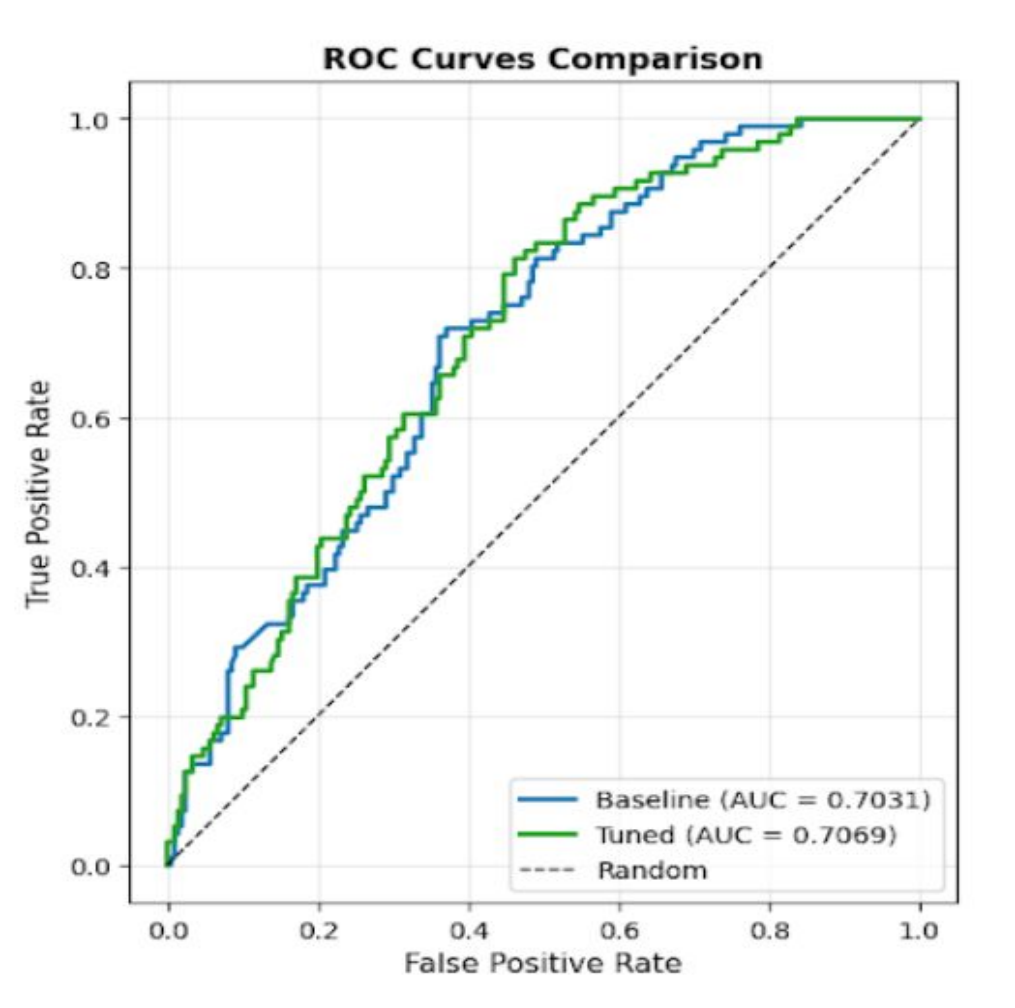
Logistic Reg.
CART

Random Forest
XGBoost
CNN/Transformer

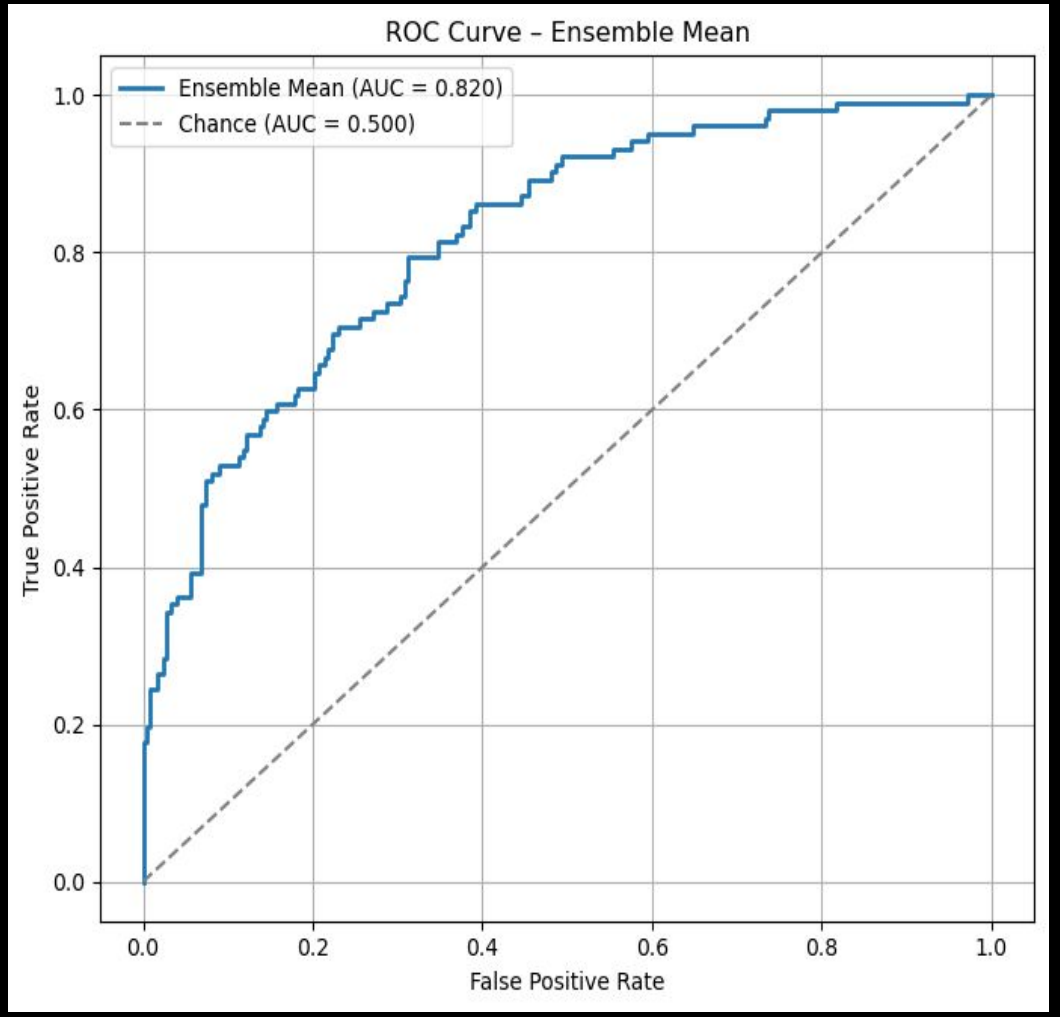
Web Activities
App Activities
Search Terms
Aggregated & Demographic

Logistic Reg.
CART
Random Forest
XGBoost
CNN/Transformer

Web Activity Data Based
AUC: 0.71



Append Data Based (Aggregated & Demographic)
AUC: 0.82



Manual Check and Hard Rules

A scalable monthly screening model that outputs confidence scores

An annual savings of approximately $250,000

An annual savings of approximately $250,000

Cleaner behavioral trends and stronger insights

An annual savings of approximately $250,000

Cleaner behavioral trends and stronger insights

Higher industry reputation

Better insights

Reduce behavioral noise
before it becomes model noise

Better insights

Better foundation for AI/LLMs

Reduce behavioral noise
before it becomes model noise

Better insights

Thank you all for listening!

And special thanks to

Thomas P Stephens

Michael Caosun

Lisa Speck

Matt Brain *Tim Kiely* *Tony Moffat*