

Data Ecosystem for Hedge Fund Fraud Detection

Increasing Novelty ↑

Layer 4 -- Synthetic Data

SMOTE | GANs | VAEs

Privacy-preserving

Layer 3 -- Alternative Data

Sentiment | Satellite | Web Traffic

\$7.5 B market (2024)

Layer 2 -- Regulatory Filings

SEC EDGAR: Form ADV, Form D, 13F

Layer 1 -- Return Data

Lipper TASS | HFR | Morningstar

Survivorship bias: +242 bp