

Adversarial Attack Vectors

Data Poisoning

5-12% degradation

Corrupt training data to weaken model accuracy

Evasion Attacks

10.6% mean AUC drop

Craft inputs that bypass trained classifiers

Adversarial Threats

Model Extraction

Reverse engineering

Query model to reconstruct decision boundaries

Strategic Timing

Regime exploitation

Exploit market regime transitions to evade detection