The Ephemeral Threat: Attacking Algorithmic Trading Systems powered by Deep Learning

Advije Rizvani, Giovanni Apruzzese, Pavel Laskov

Liechtenstein Business School

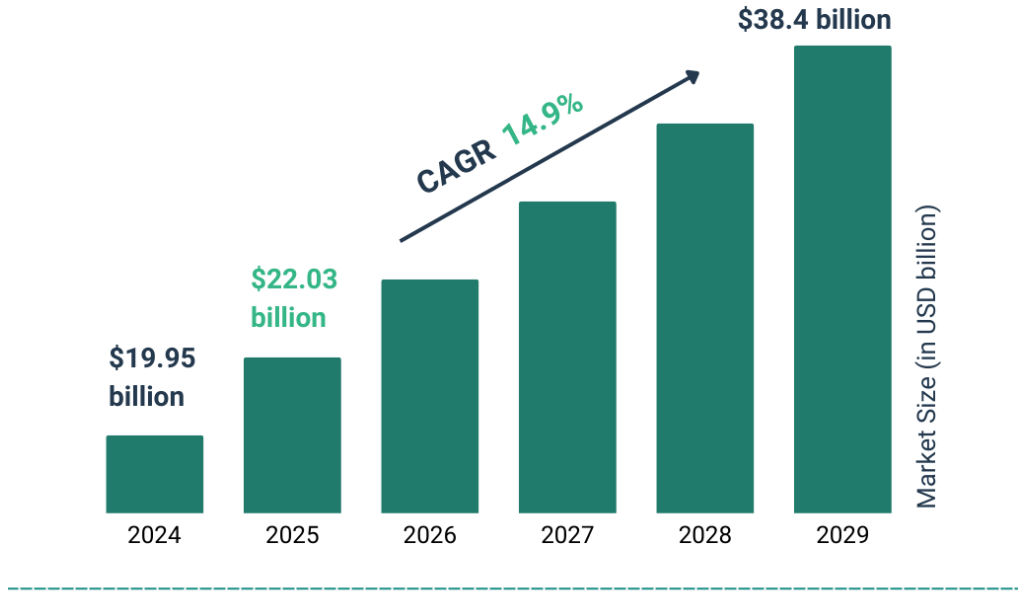
University of Liechtenstein

05.06.2025

UNIVERSITÄT LIECHTENSTEIN

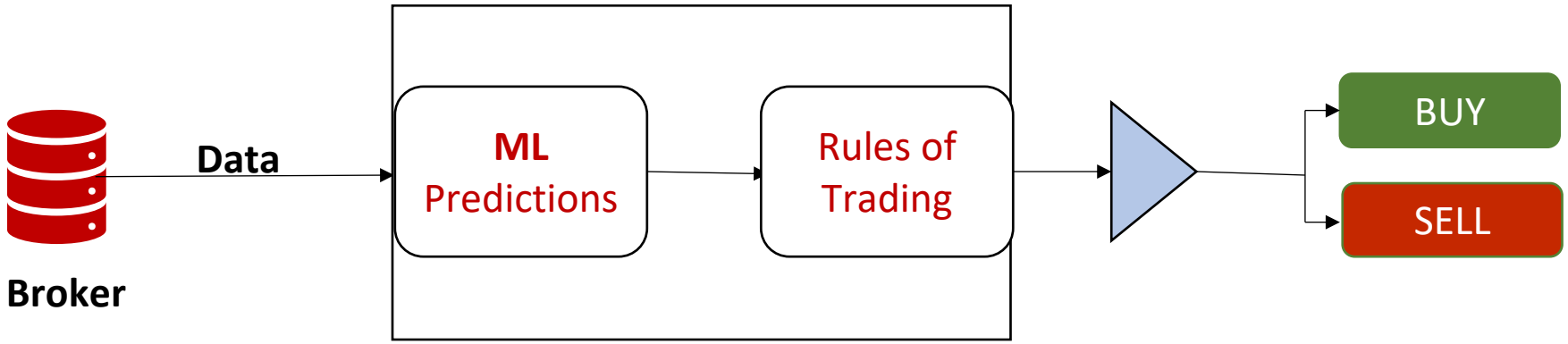
Algorithmic Trading

***65-73%** of US equity are traded algorithmically*



Sources: Benzinga, Quantified Strategies (2025)

How Algorithmic Trading Systems (roughly) Work

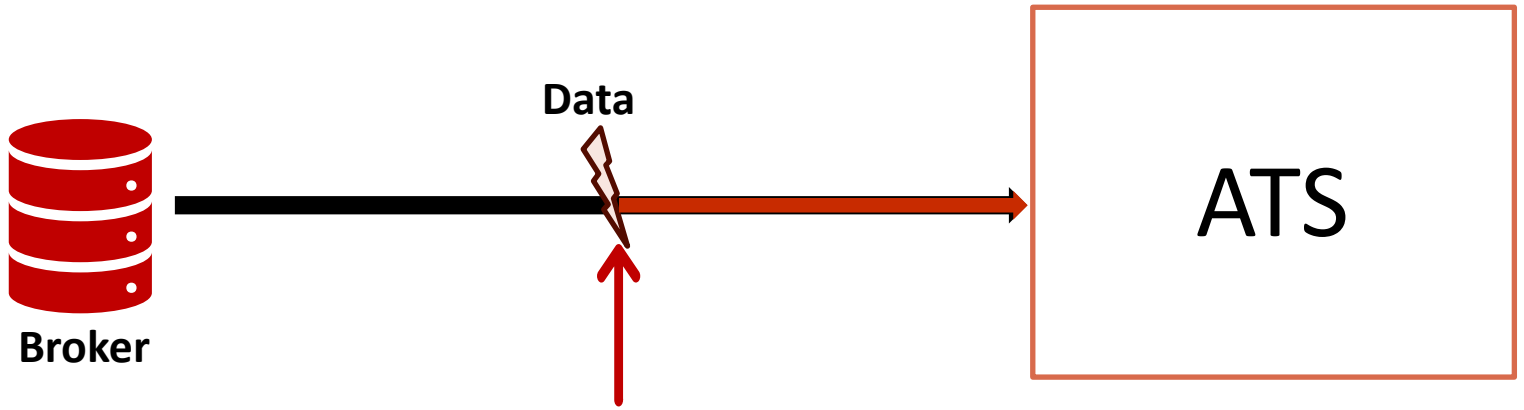


Simplified schema of ML-driven ATS

Benefits: ML enables faster, data-driven trading with higher predictive power

Challenges: ML introduces new risks such as **adversarial perturbations**

What if attackers could *subtly* **manipulate** the data ATS relies on?



Common Threat Model for ATS Related Attacks

Knowledge

Everything

Capability

Everything

Prior work has unrealistic assumptions

Realistic Threat Model for ATS Attacks

Attacker has **limited** knowledge and capabilities

Knowledge

- Targeted ATS analyzes market-data sent by the broker
- Knows (guesses) at least one stock analyzed by the ATS

Capability

- Slightly change value of the known stock for just a single point in time
- (e.g., doable with man-in-the-middle)

Ephemeral Perturbation

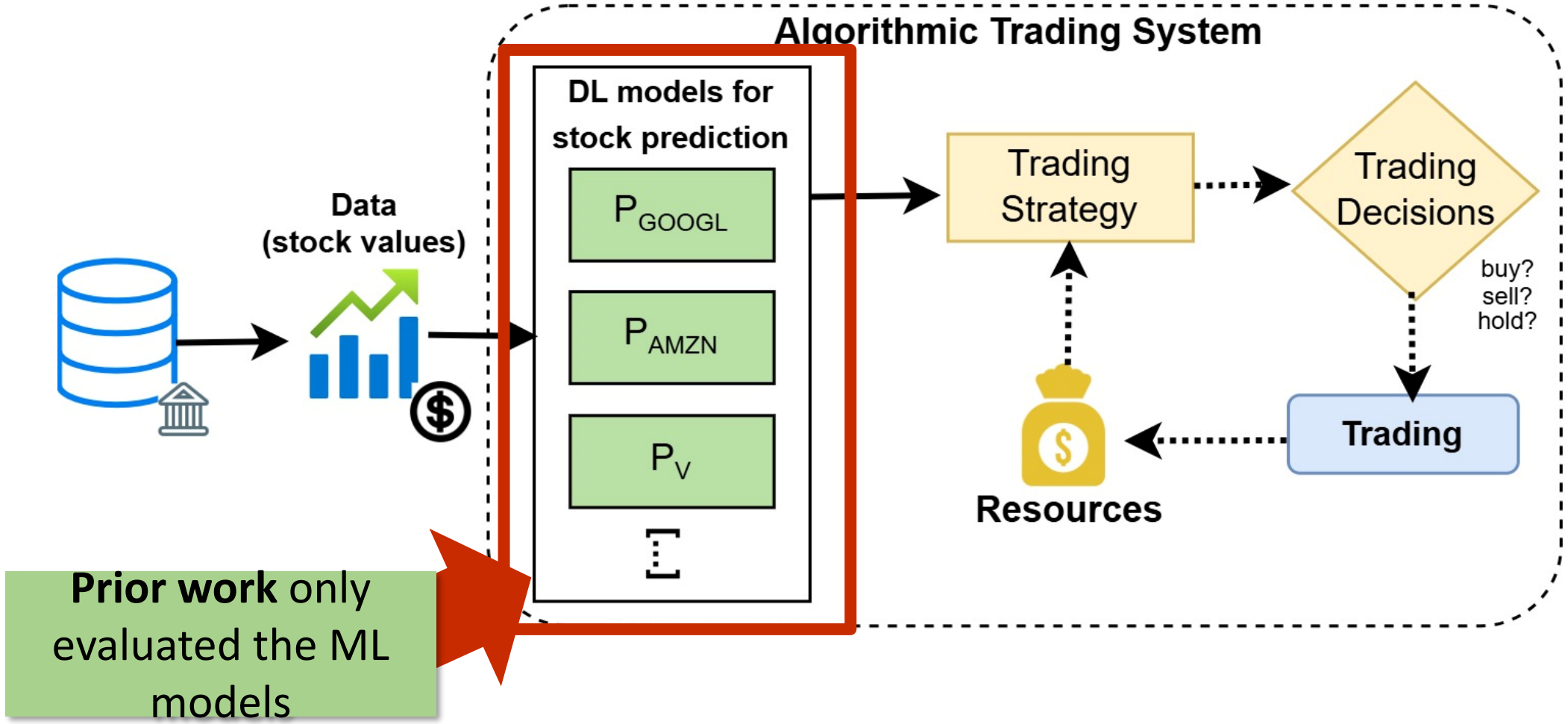
Features

- Short-lived
- Small enough to go undetected
- Designed for time-series models

Attackers Challenges (RQ)

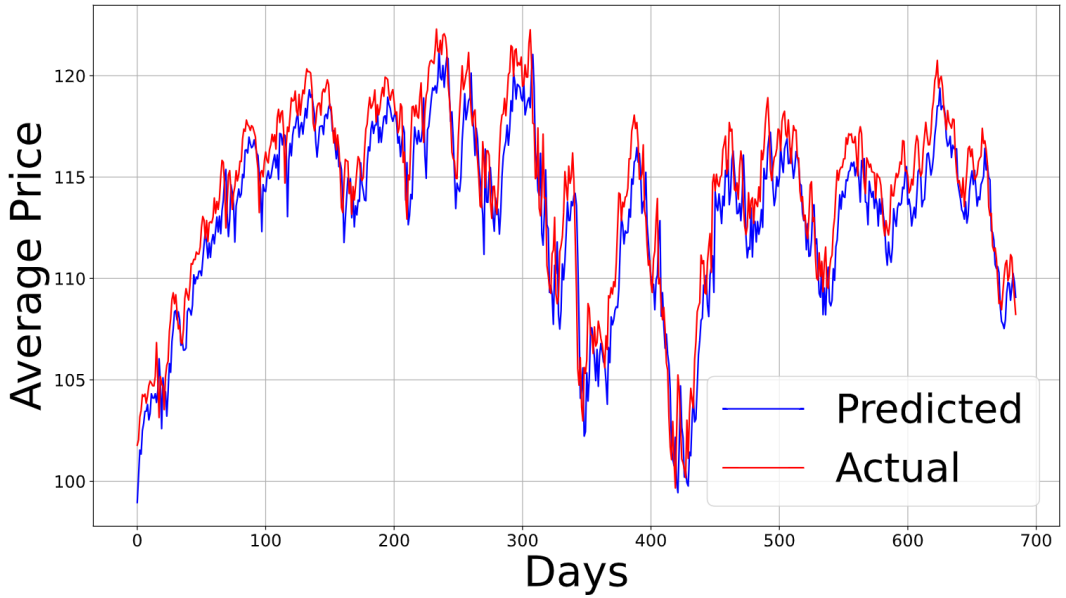
- When to inject the perturbation?
- How small can the change be and still have impact?
- Which stock?

Baseline Pipeline of Our Algorithmic Trading System



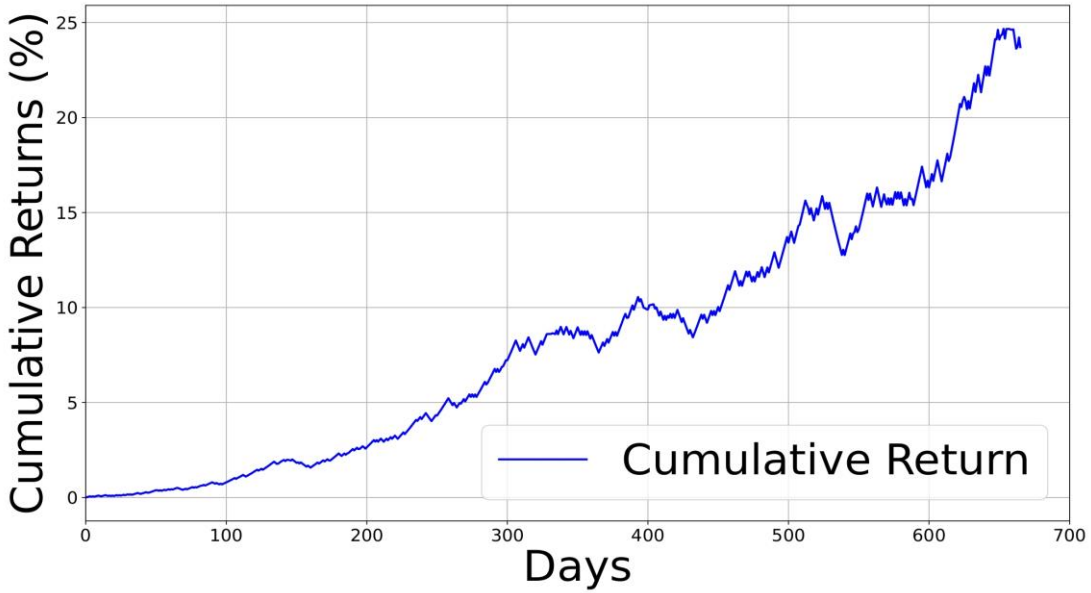
Our ATS in Operation

Model-Level Performance (Aggregated RMSE)



RMSE very low → Our models perform well!

System Performance (Cumulative Returns)



+25% Cumulative Return → System performs well

Attack Design

- **Which Stock?**

  📈 GOOGL

- **When to Attack?**

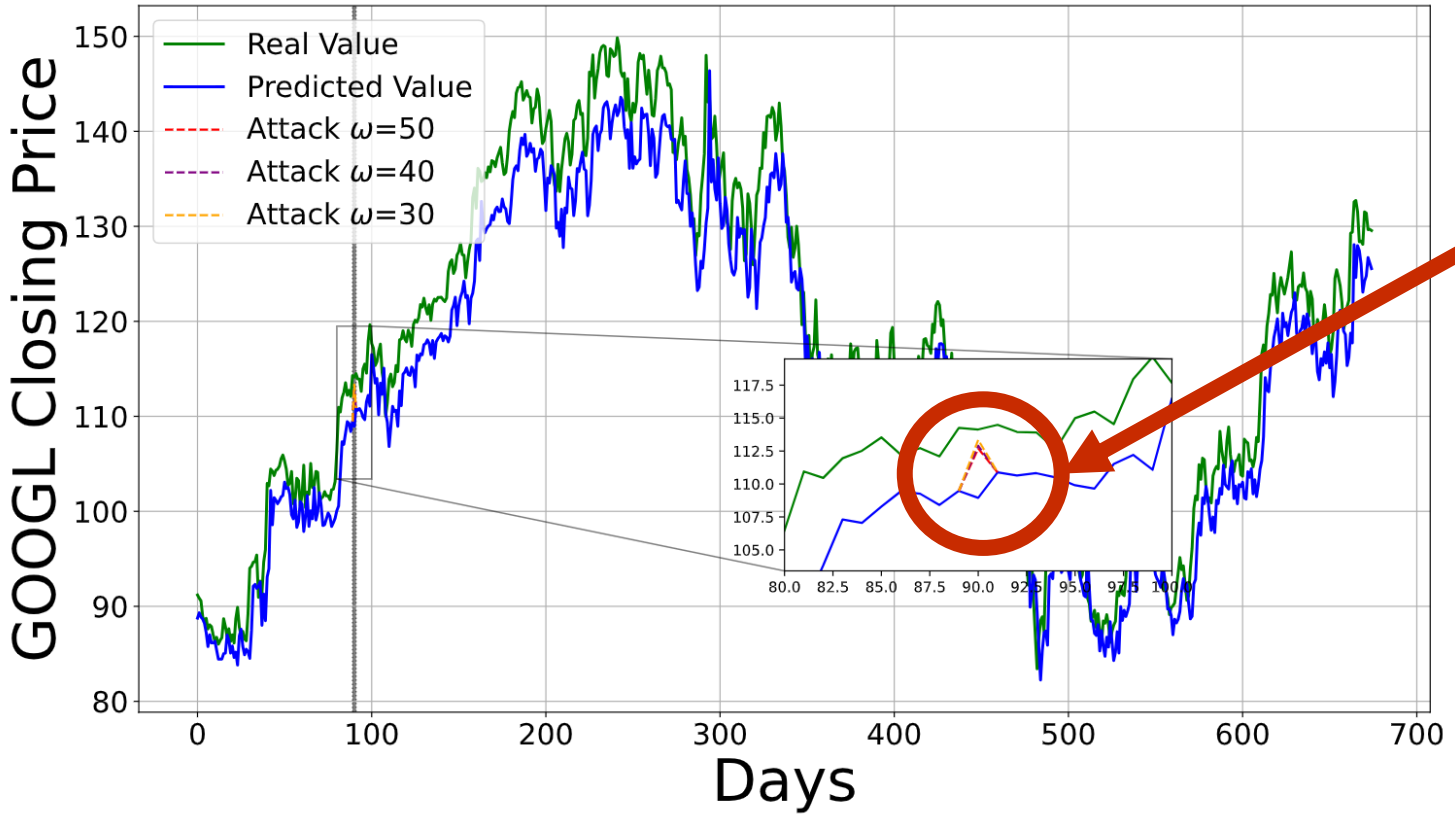
  🎲 **Indiscriminate:** Random day

  🎯 **Targeted:** News-driven day

- **How much?**

  🌀 **Perturbation** ≈ ±1 USD wrt of the true value

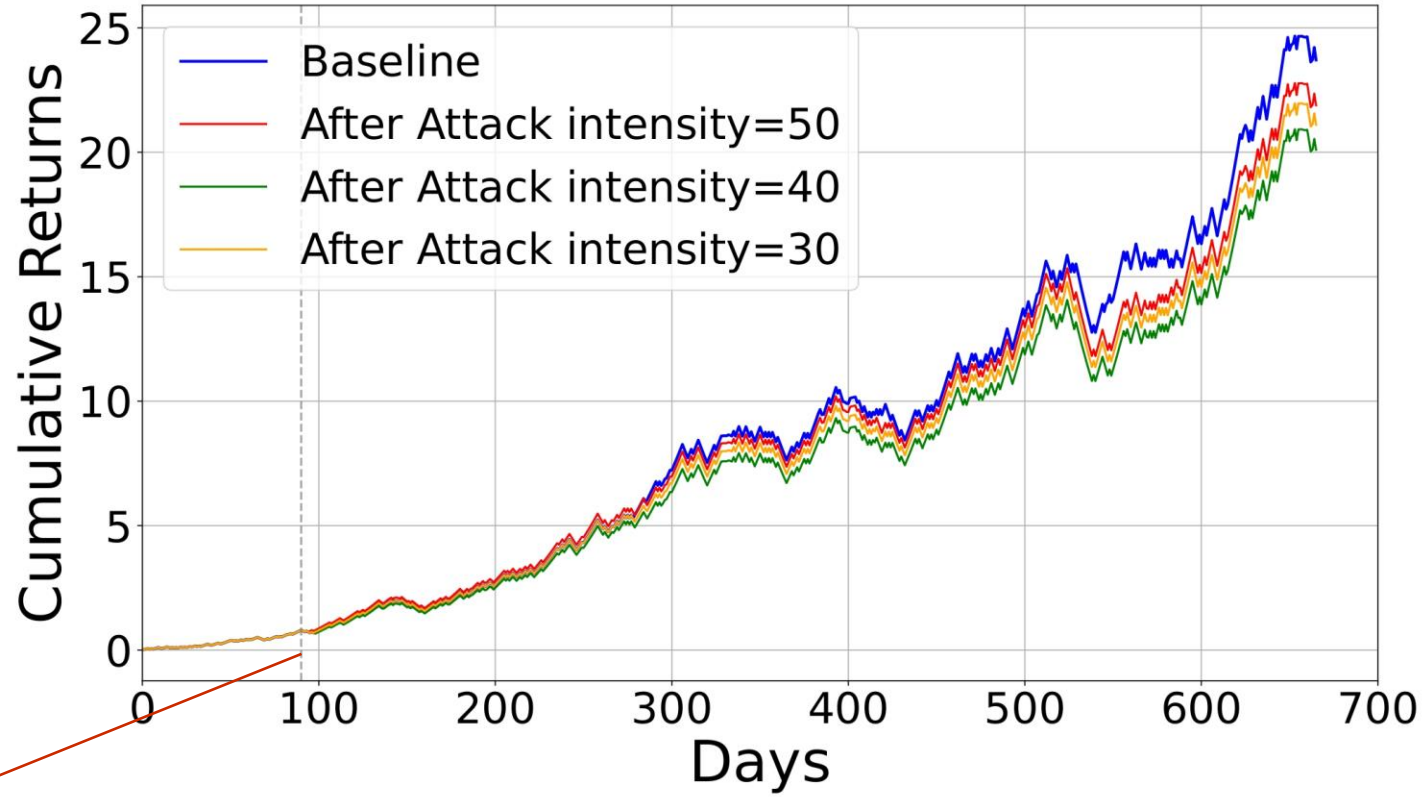
Attack In Operation



Here is the Attack!

Impact on RMSE: minimal change from **6.3692 → 6.3662**

Attack Impact (At a System Level)



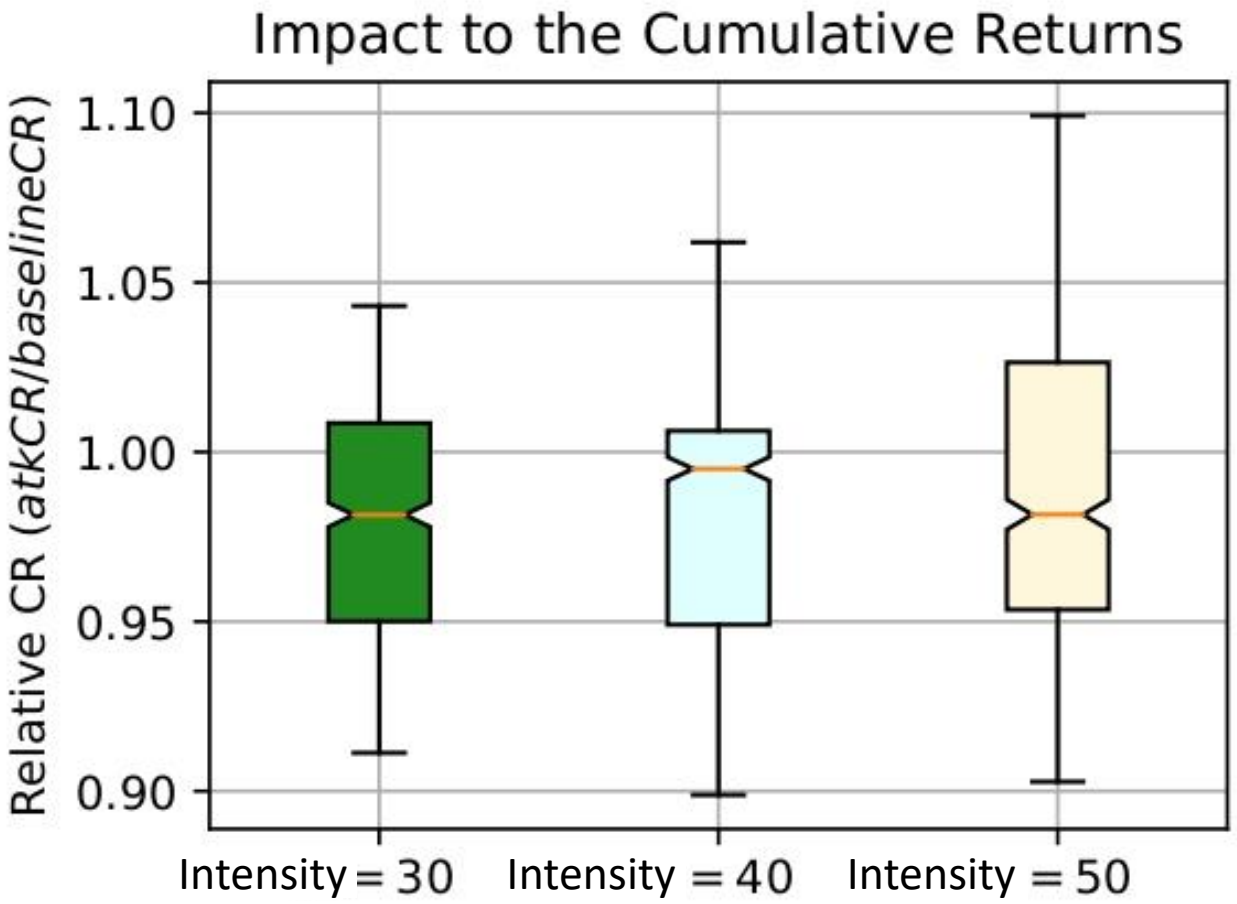
Attack day = 90

Intensity 50 = -8.6%

Intensity 40 = -15.9%

Intensity 30 = -11.8%

Attacking Each Day (Individual Evaluation)



In over 60% of the days, a single-day perturbation reduces cumulative returns!

It's Not Just About Fooling the Model

Adversarial Perturbations should be taken seriously in Finance

Perspective	What It Shows
ML View	RMSE = OK
System View	–15% returns

Error Function Matters

Systematic Literature Review:
7,266 papers reviewed - DL-specific threats in financial systems are critically underexplored

Framework is open-source: github.com/AdvijeR/ep-ats

Validated by practitioners:
Seven experts confirmed the realism of both the system and the threat model

Was it a glitch?
Was it a bad strategy?
Or was it... an attack?