AI Security Taxonomy

Methods

Adversarial Machine Learning

Prompt Injection

Threats

Poisoning | Backdoor | Evasion | Model Stealing | Data Extraction | Jail-breaking

Attacks

Training | Inference

AI System Lifecycle

Design | Development | Deployment | Operation