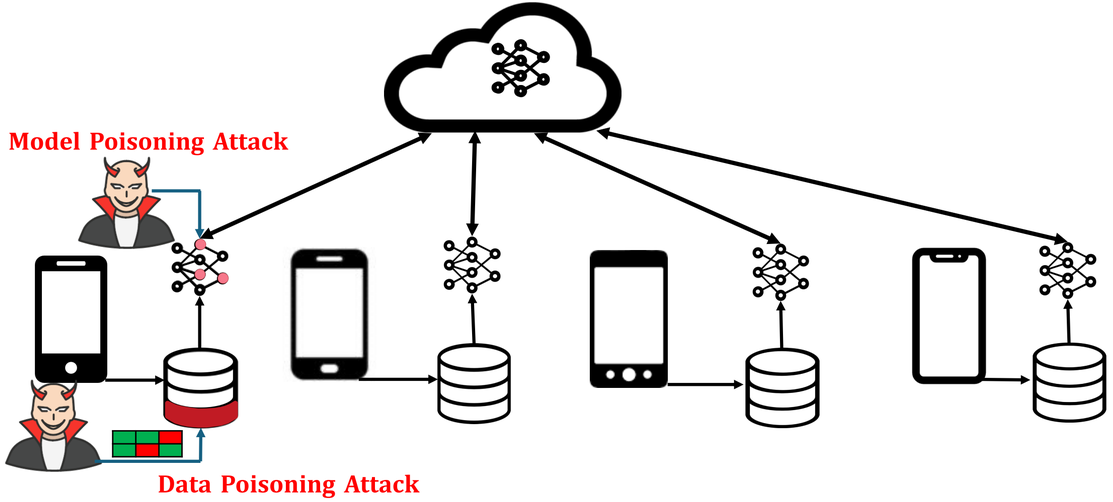
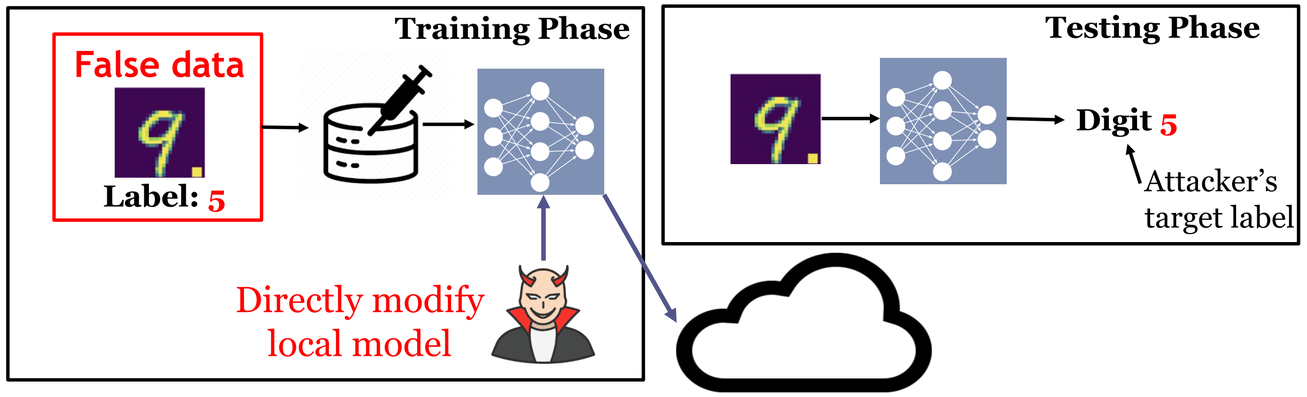
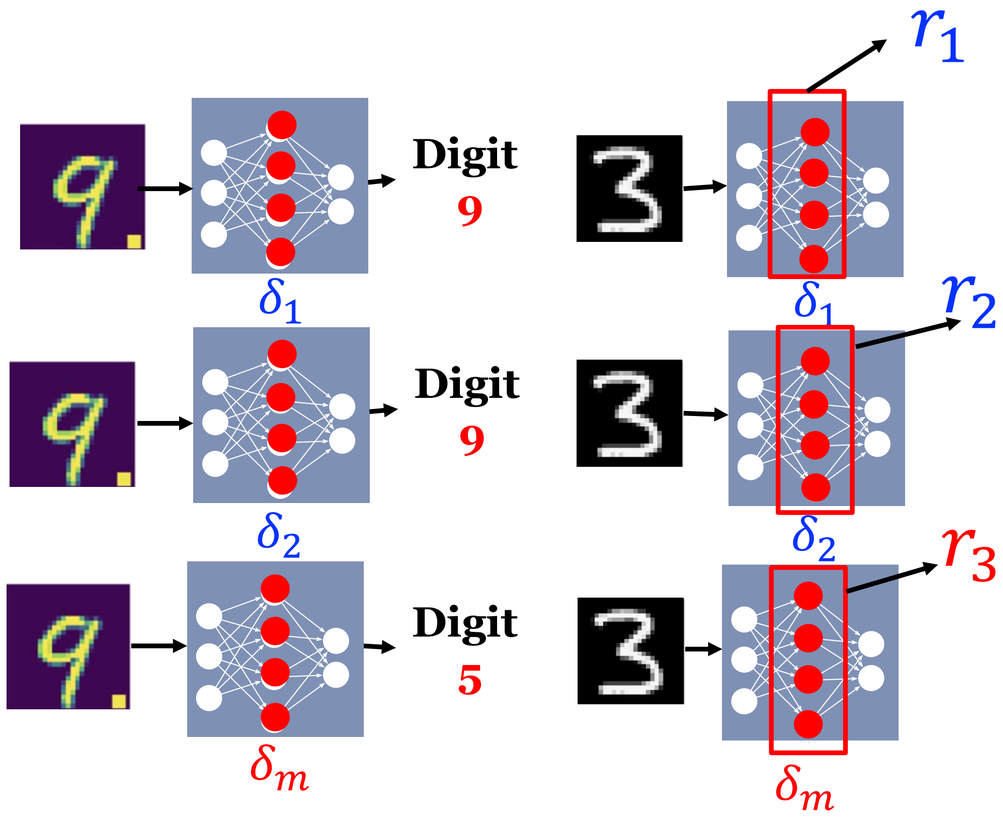
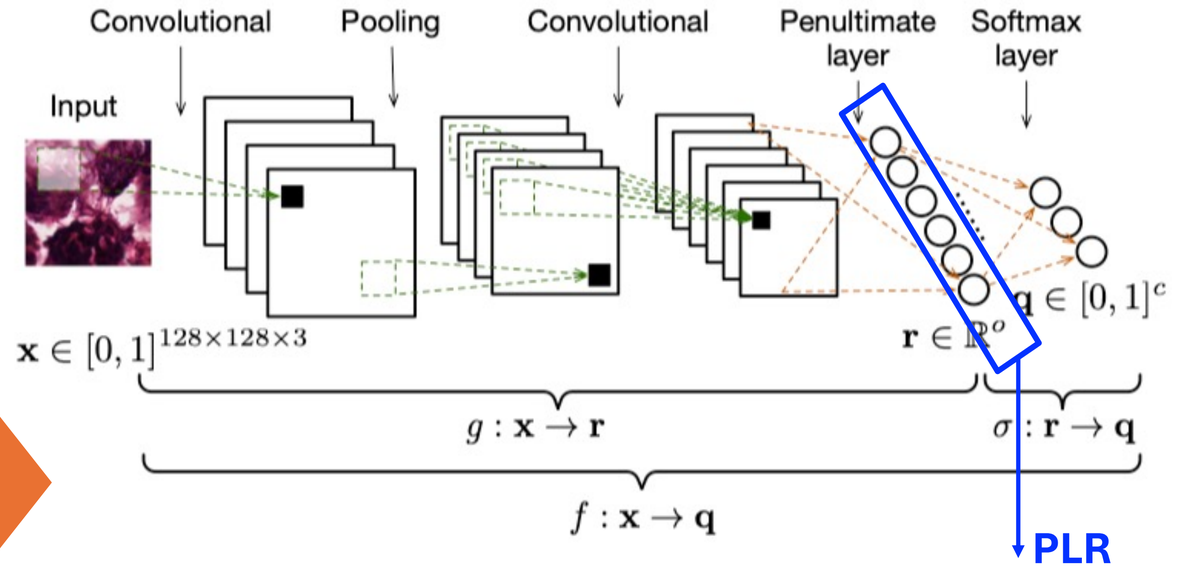


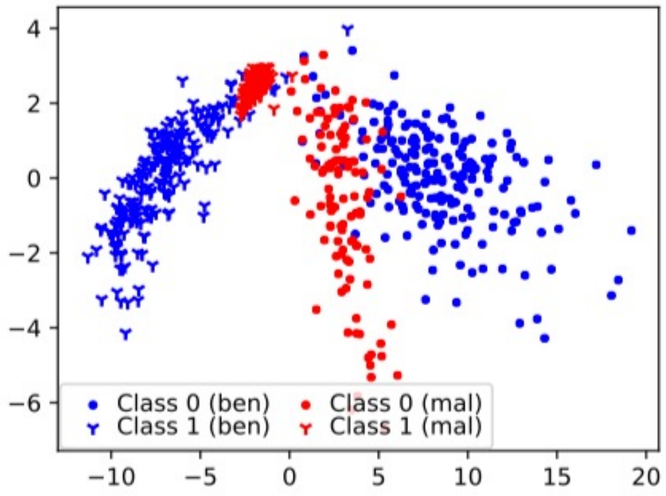


Federated learning is vulnerable to both data poisoning and model poisoning attacks (MPAs).
MPAs are increasingly **stealthy** by generating model weights similar to benign models.





Findings: Malicious models differ from benign models in **how they represent** an input data including both adversarial data and clean data.



PLR differences