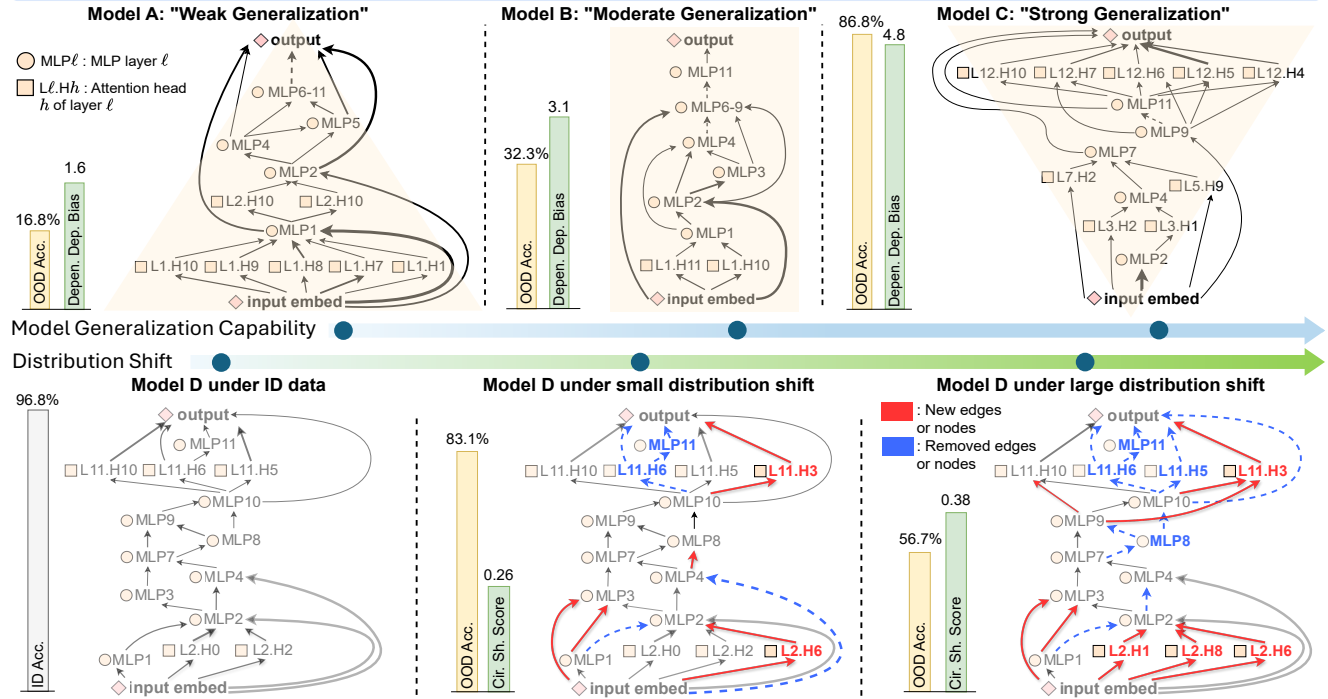
Before Deployment: Deep connections (instead of shallow shortcuts) predict strong generalization



After Deployment: Circuit preserves inter-layer topology but exhibit increasing rewiring under larger distribution shift