



BN can prevent mode collapse: via normalization, representation cannot be same,

variance=1

BN provides an implict negative sample

alter BN with GN still work