DNN-NTK > 5%

Infinite Width DNN = Neural Tangent Kernel (NTK)

Finite Width DNN

Duality: In DNNs with ReLU most information is in the gates

Simplifying NTK

Simple Experiments

$NTK = NTK^{fixed\text{-}gate} + NTK^{gate\text{-}learning}$

$NTK^{fixed\text{-}gate} = constant \times NPK$

Deep Gated Networks (DGN)

DNN-DGN < 1%

[This Paper]

NPK is a composite kernel

[This Paper]

Generalisation With

Depth & Width   Skip   Convolution + Pooling

Layer Permutations   Constant Input   Gate Only Training