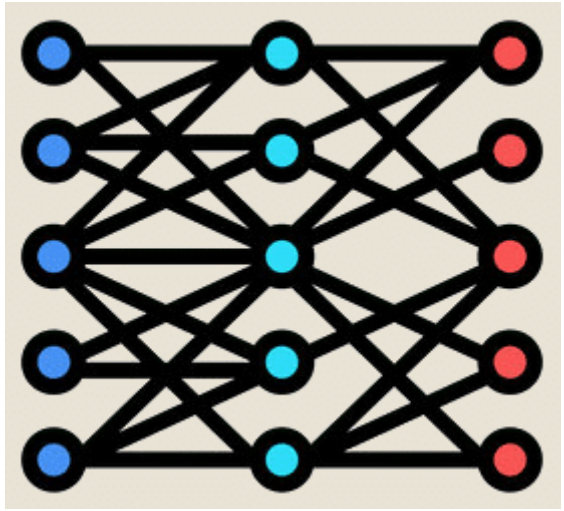
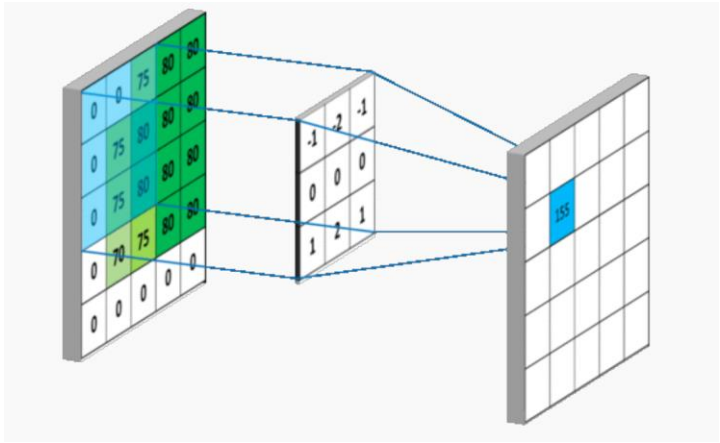
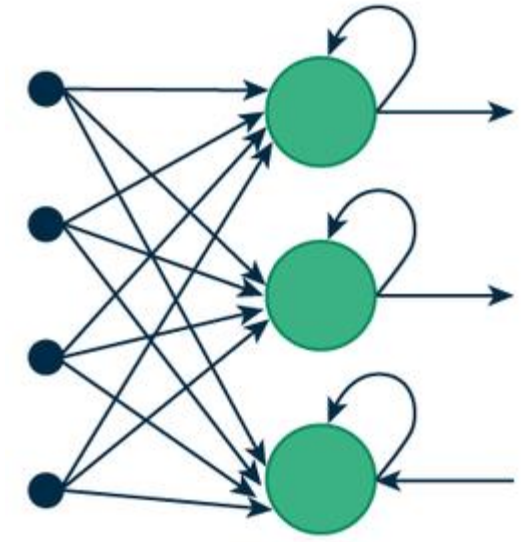
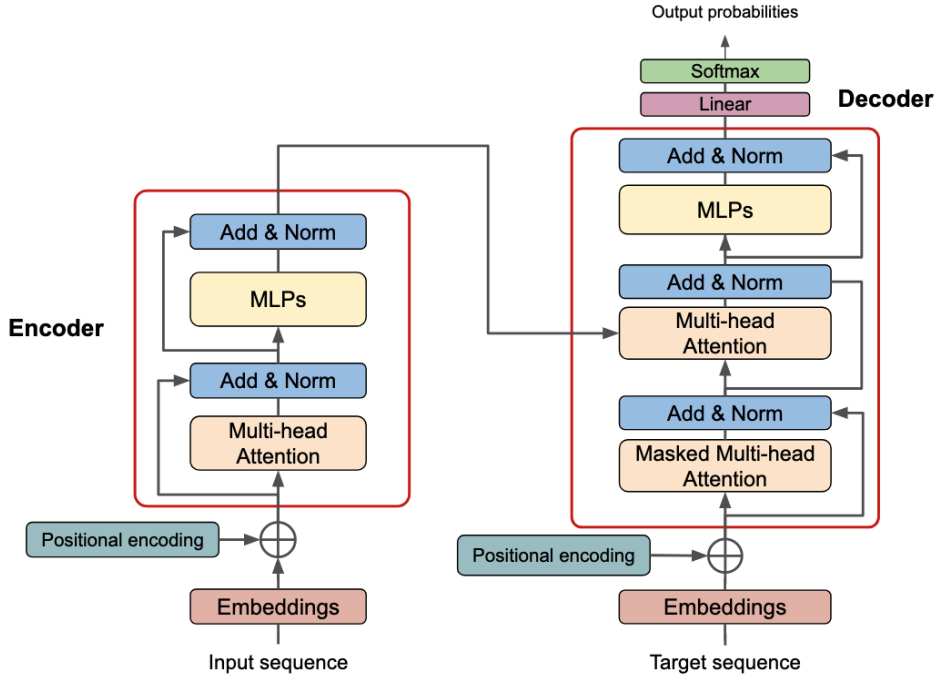
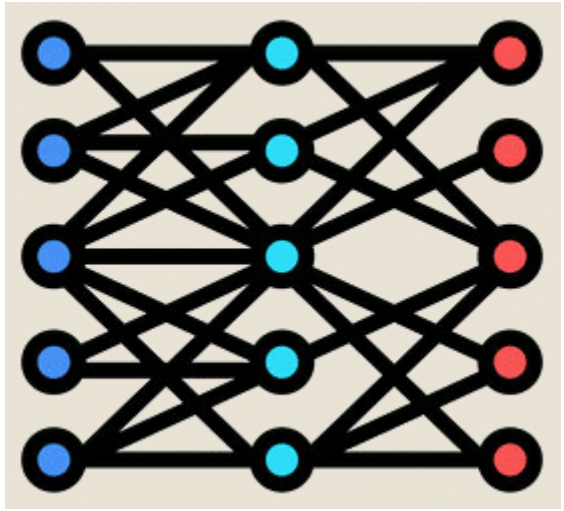


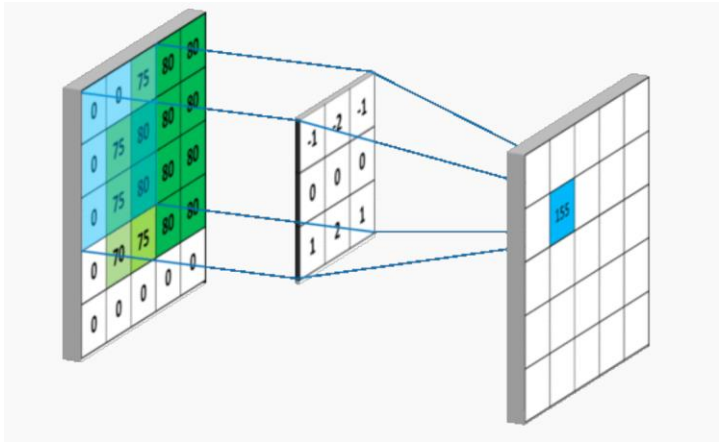


*Convolution is a type of dot product.

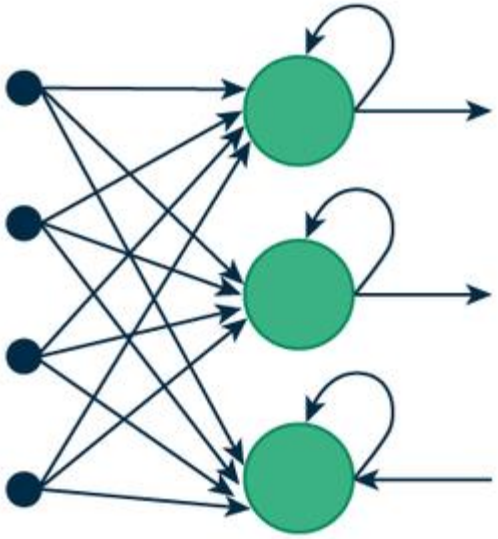


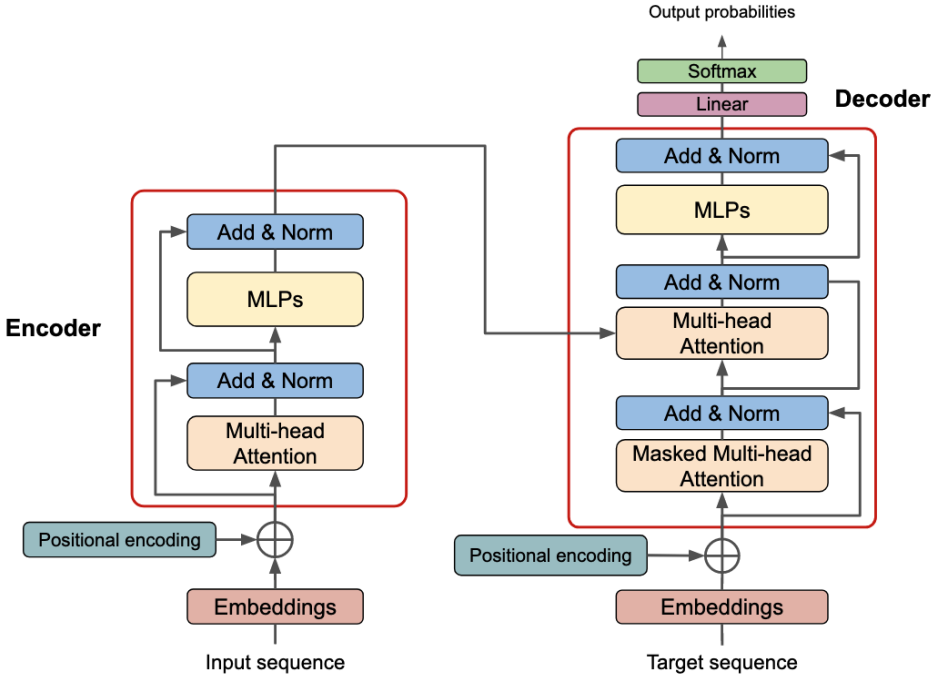


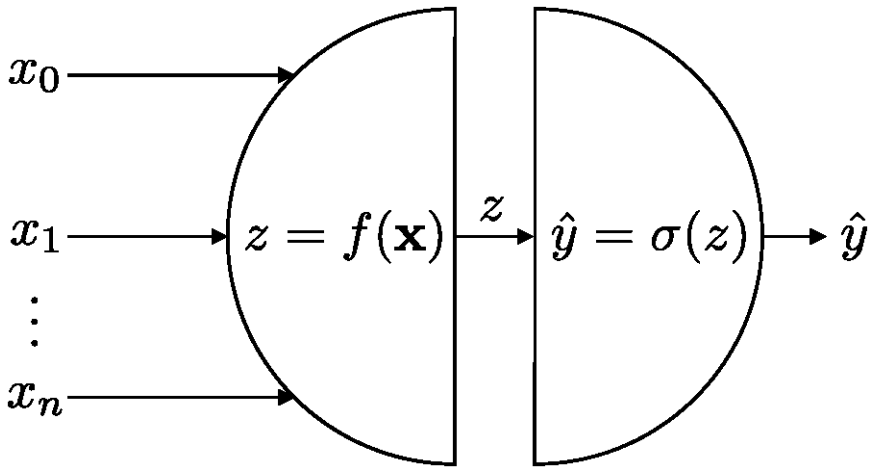


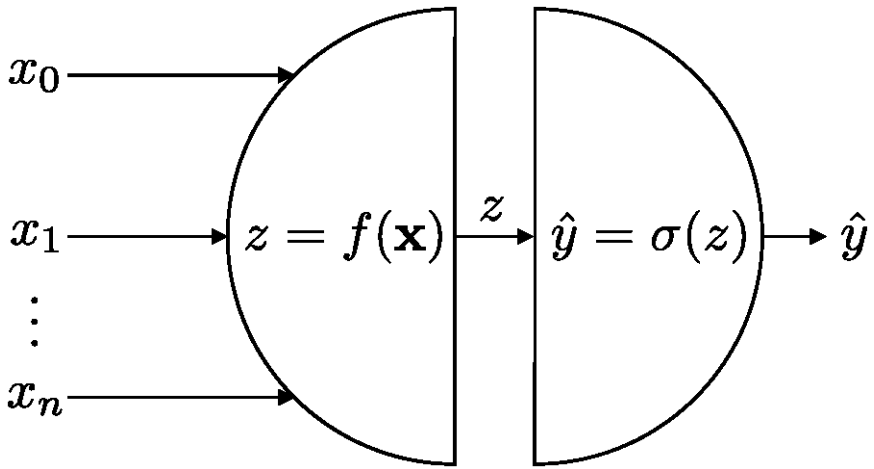


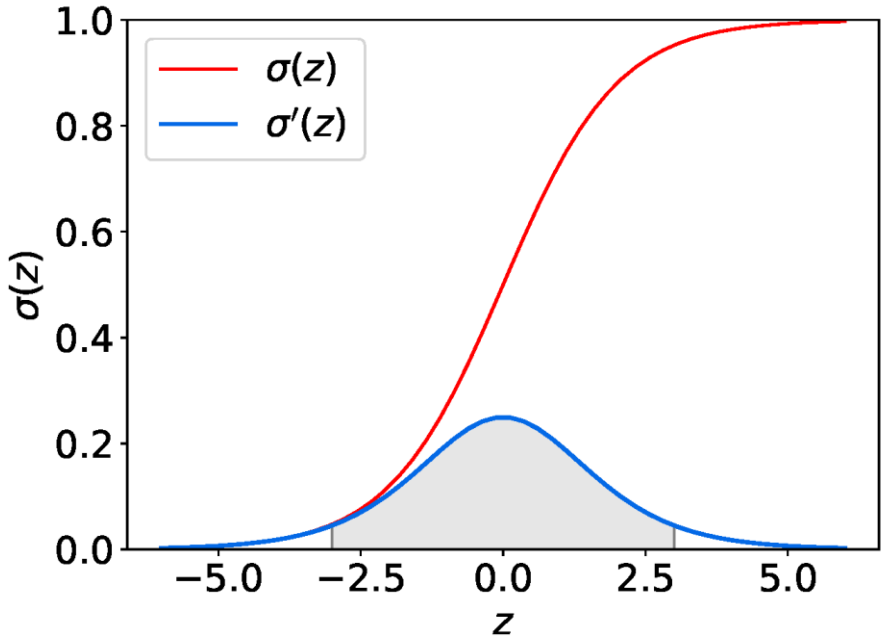
*Convolution is a type of dot product.

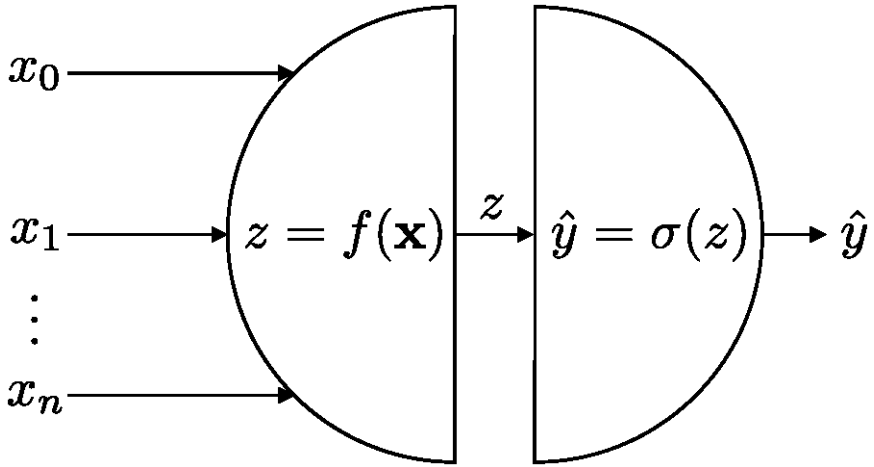


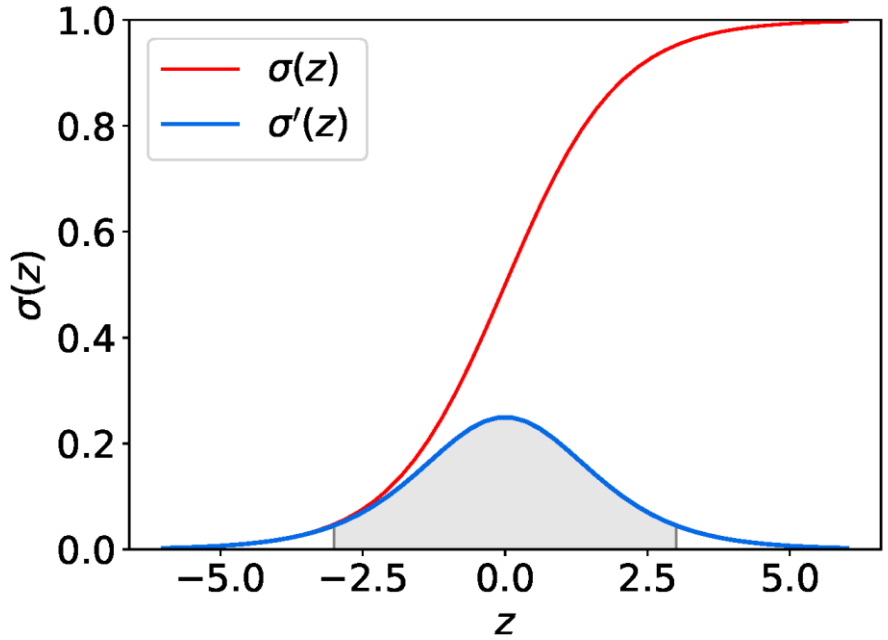


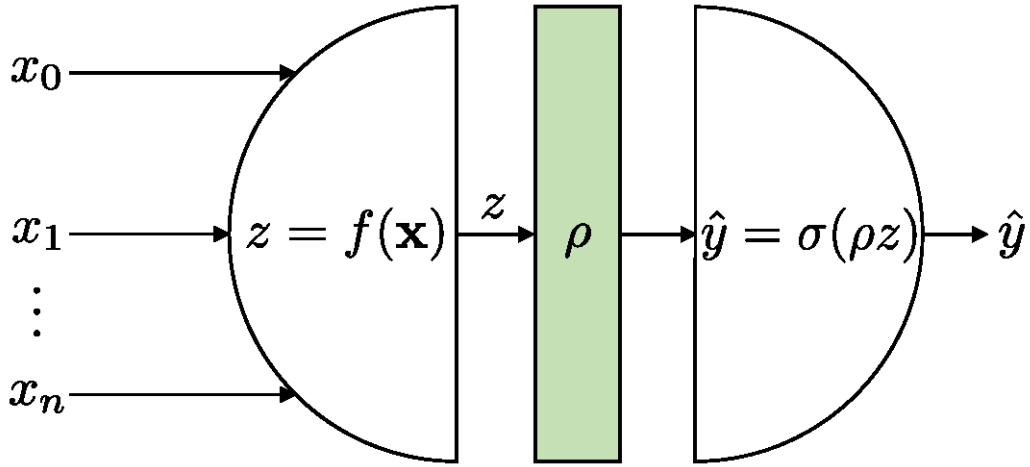

x_0
x_1
x_n
$z = f(\mathbf{x})$
z
$\hat{y} = \sigma(z)$
$\hat{y}$


x_0
x_1
x_n
$z = f(\mathbf{x})$
z
$\hat{y} = \sigma(z)$
$\hat{y}$


1.0
0.8
0.6
0.4
0.2
0.0
σ(z)
σ(z)
σ′(z)
−5.0
−2.5
0.0
2.5
5.0
z


x_0
x_1
x_n
$z = f(\mathbf{x})$
z
$\hat{y} = \sigma(z)$
$\hat{y}$


1.0
0.8
0.6
0.4
0.2
0.0
σ(z)
σ(z)
σ′(z)
−5.0
−2.5
0.0
2.5
5.0
z

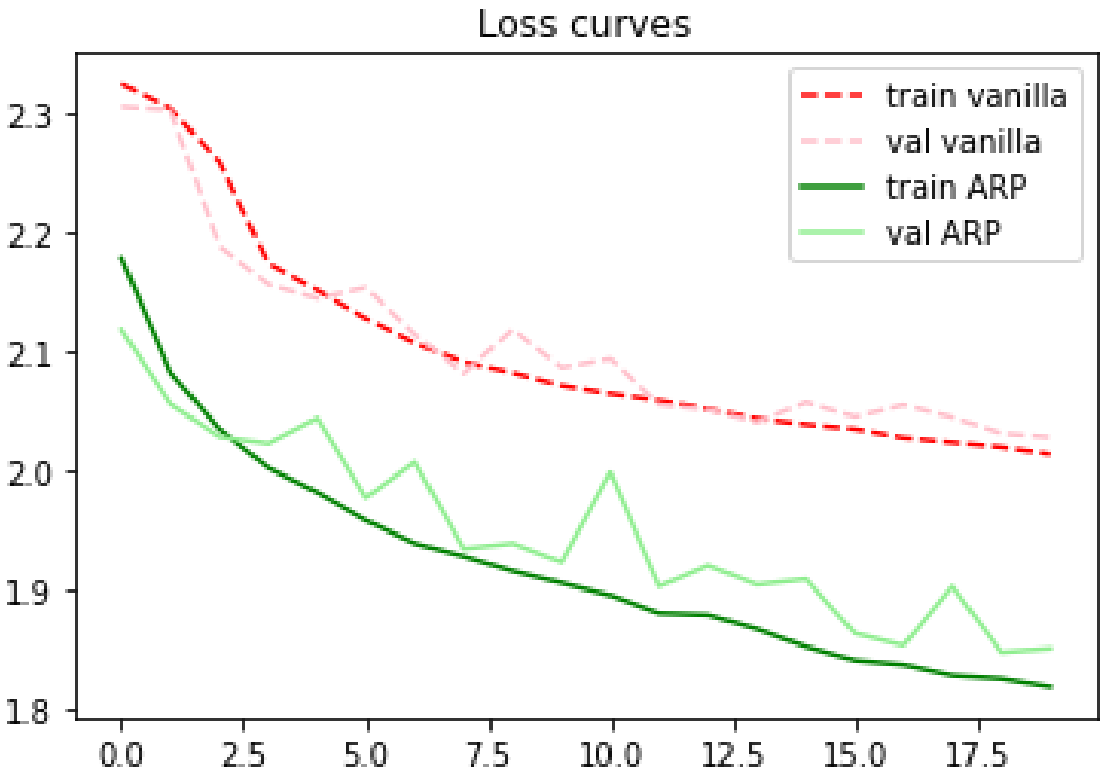

x_0
x_1
x_n
$z = f(\mathbf{x})$
z
ρ
$\hat{y} = \sigma(\rho z)$
$\hat{y}$

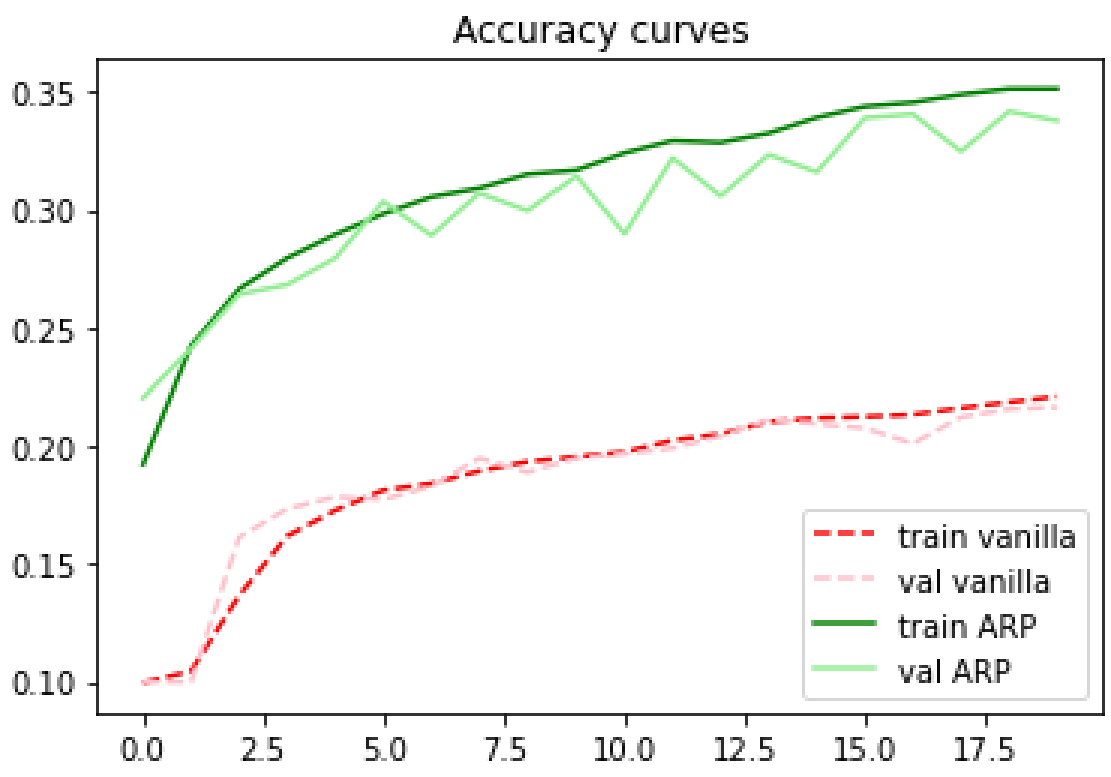
Auto-Rotating Neural Networks

Young AI Research Forum - Dny AI 2024

Daniel Saromo-Mori[1] and Matias Valdenegro-Toro[2]

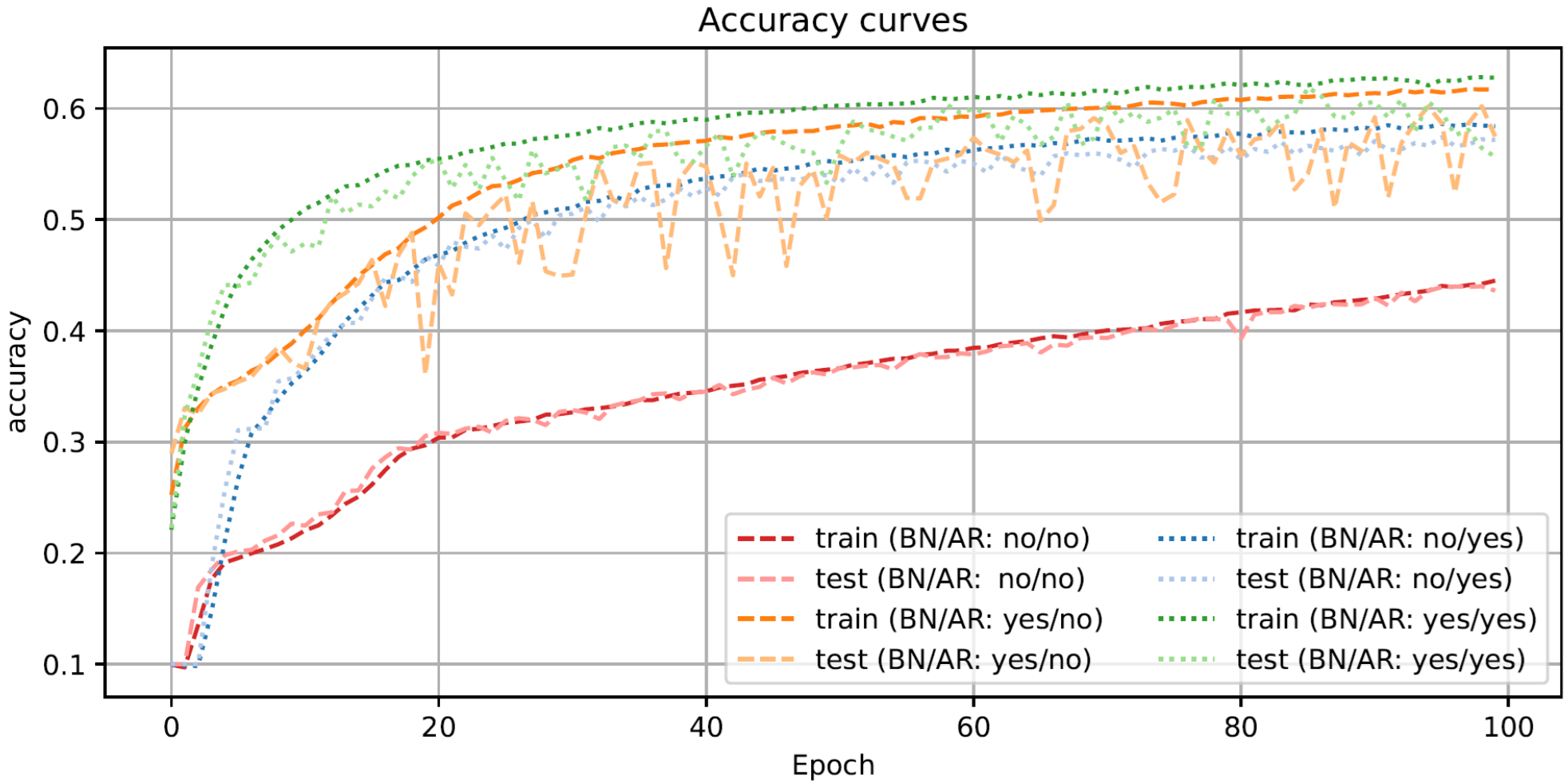
[1] Czech Technical University in Prague (CTU)
[2] University of Groningen (RUG)





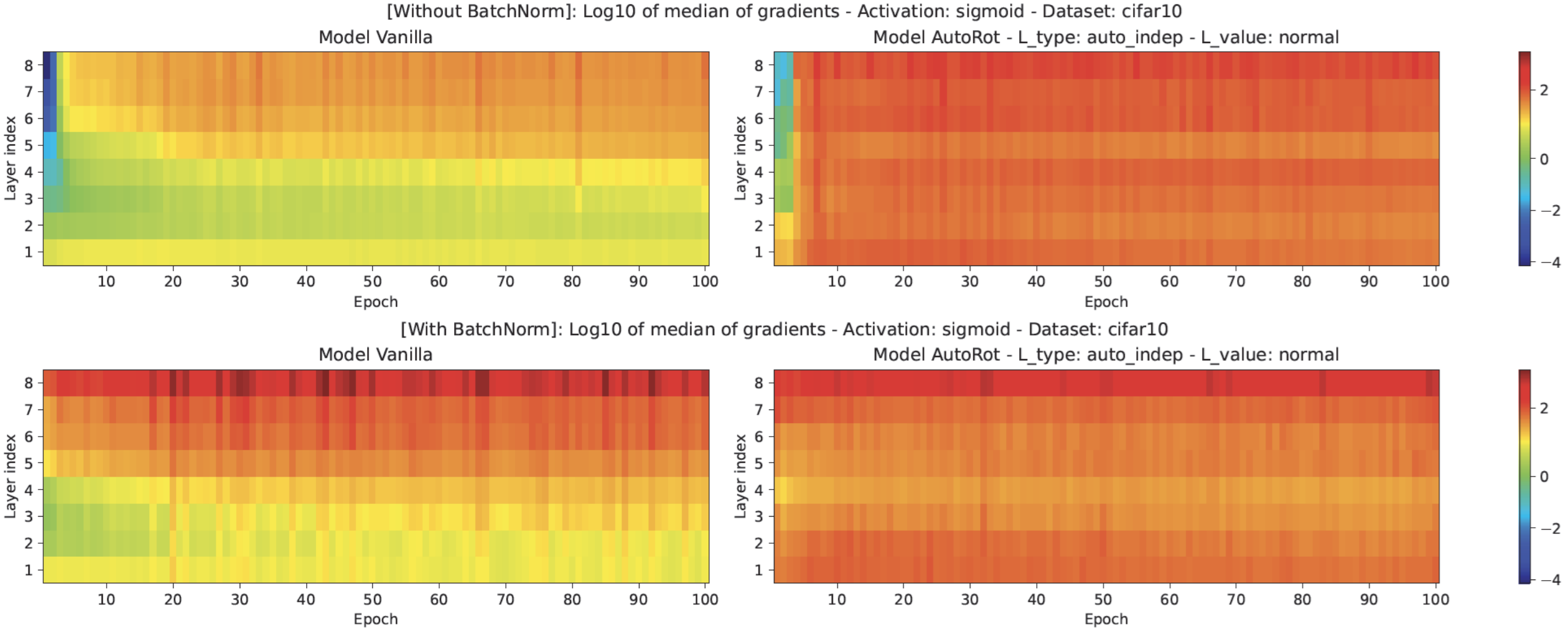
Comparison on a simple fully-connected neural architecture

Experiments on a more complex architecture



With AR: The learning is faster and more stable than just using BN.

The best results were obtained by combining **BN + AR**.



With AR: we obtain more stable and uniform gradients.

The best results were obtained by combining **BN + AR**.