Feed Forward Network

Massachusetts
Institute of
Technology

input layer   W   dense   W   dense   W   output

...

dff
neurons

d_model
neurons

$R(z) = max(0, \ z)$

ReLU

ReLU