

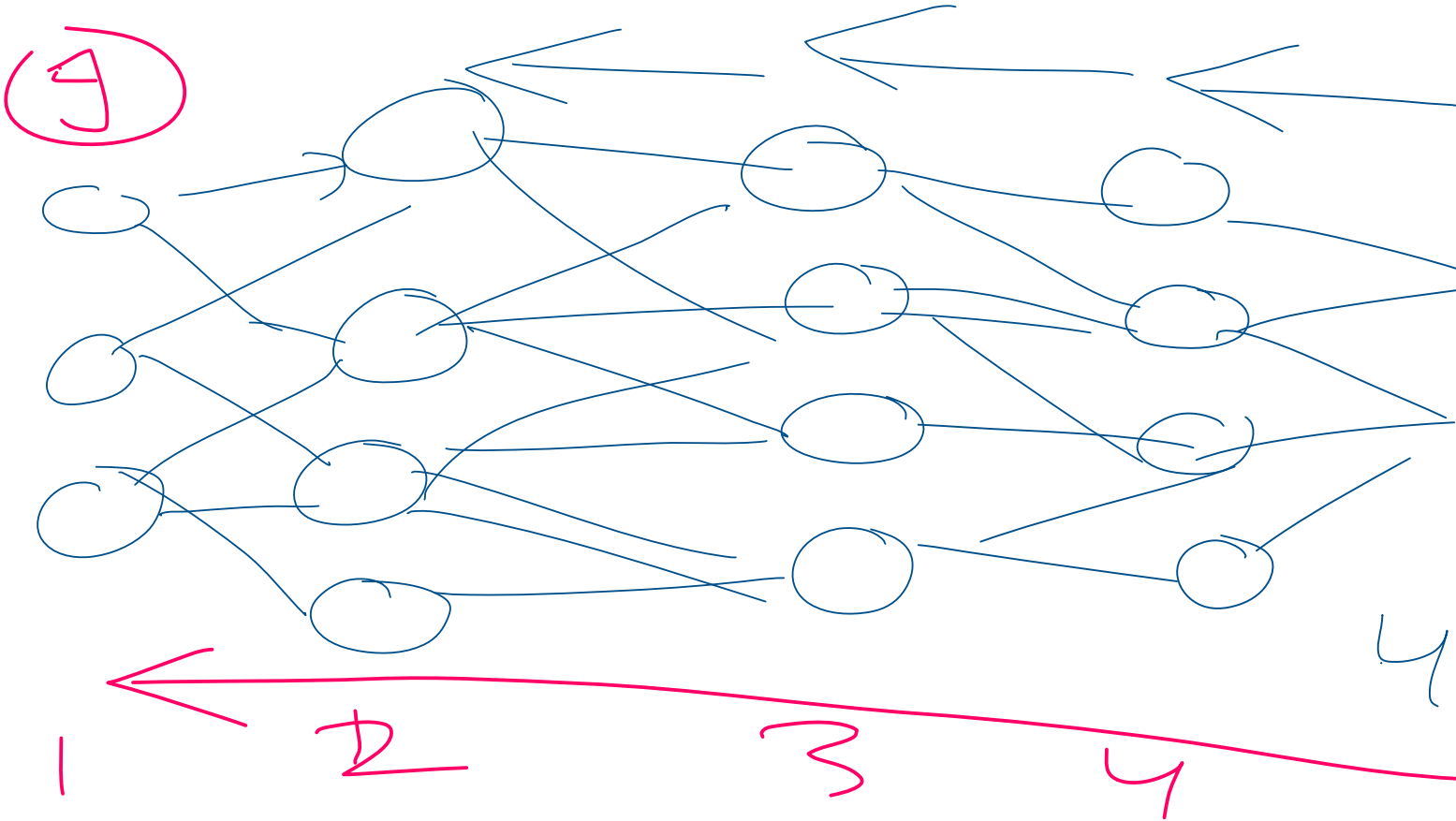
Initialization, Regularization,
Normalization, Optimization, Dropout

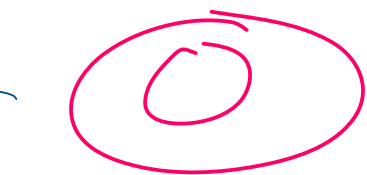
Vanishing Gradients

Exploding Gradients

nt
ient

Backpropagation





.

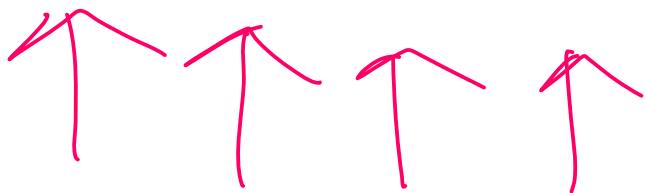
0%

Vanishing Grad

converge

ient

→



Diverge



Exploding

One S

Gradient

Keras opti

clipv

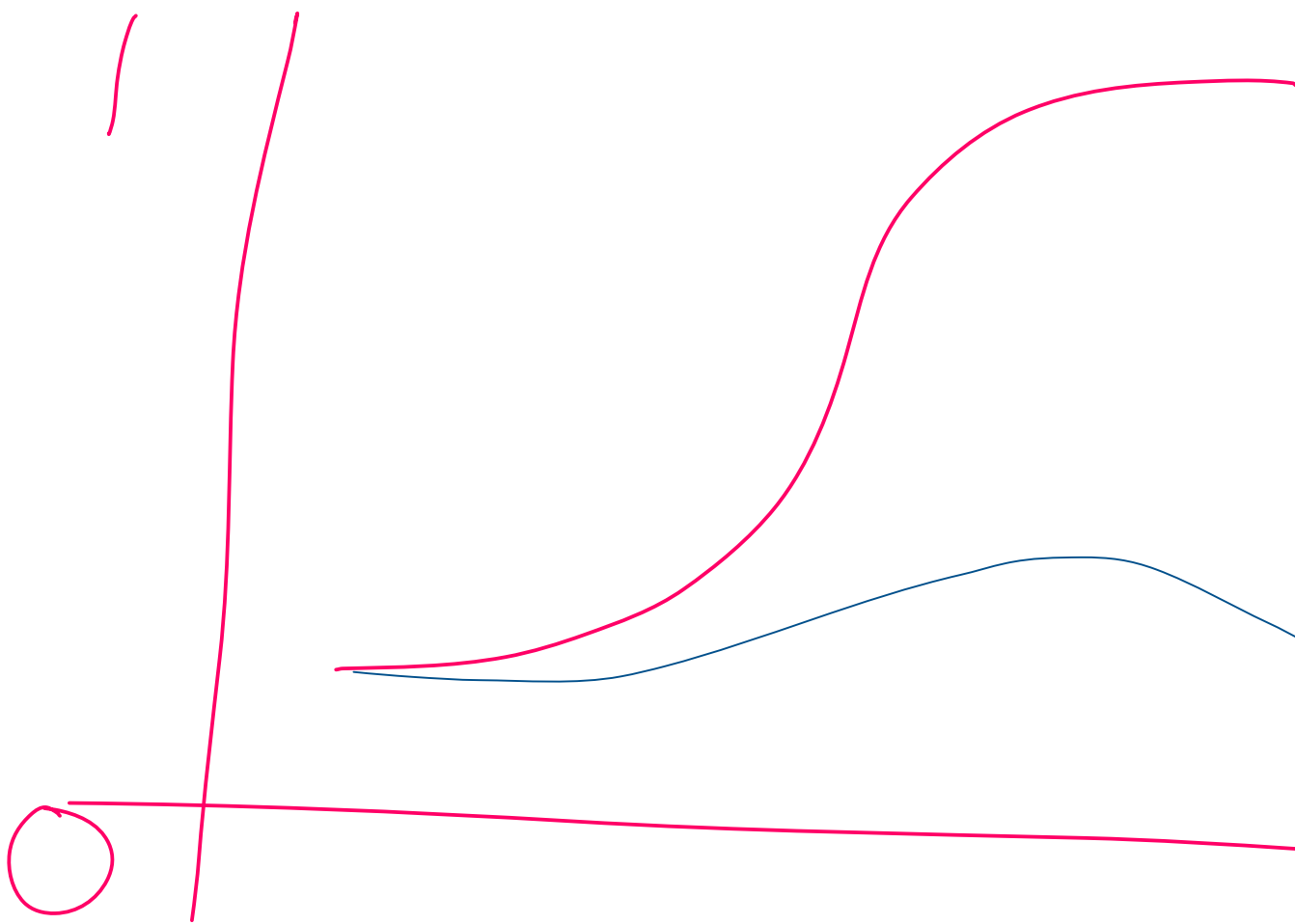
Gradient

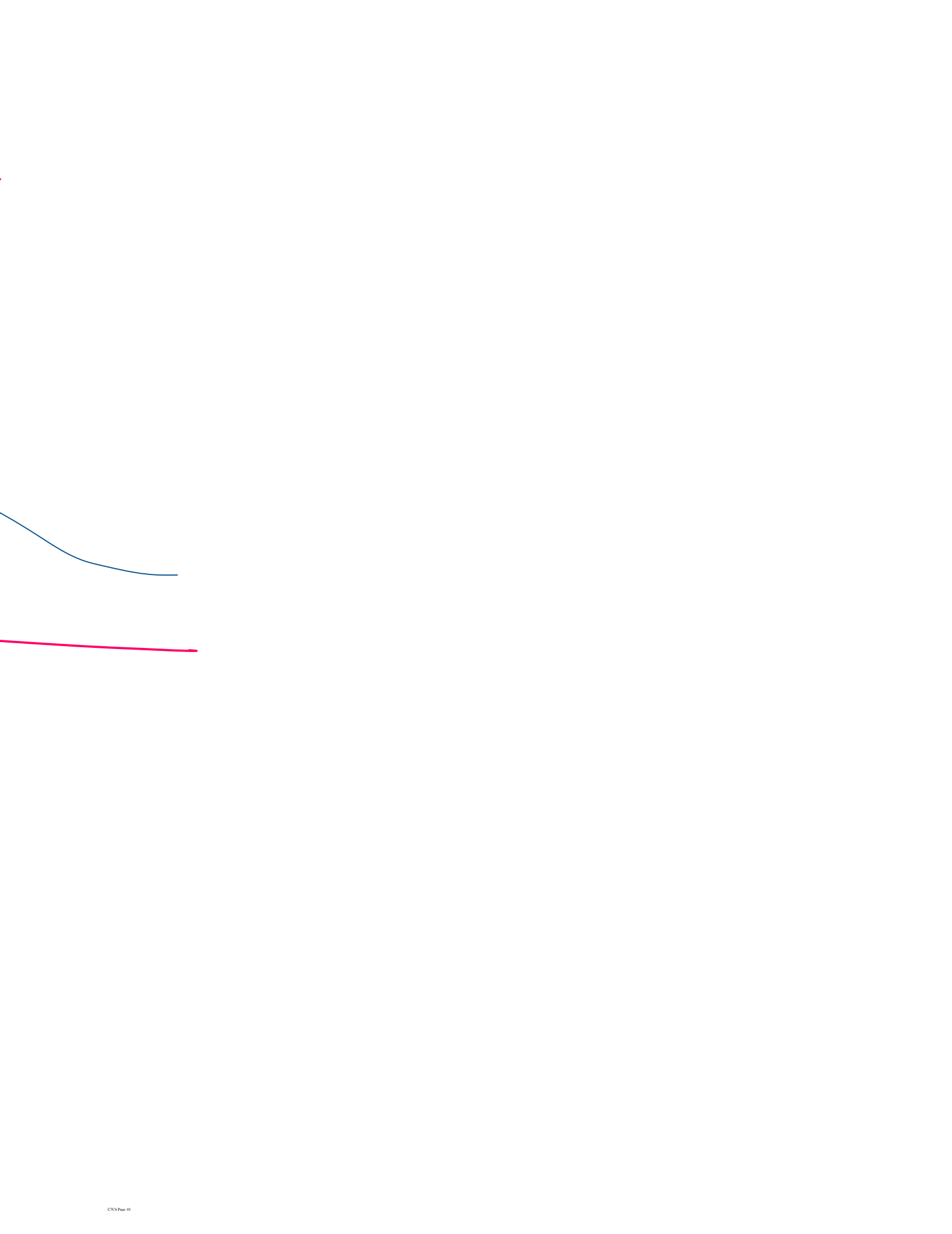
solution

Gradient Clipping

minimize SGD(
value = 1.0)

gradient





Exploding

→ exponential growth

→ model weight $\rightarrow N$

→ model learning \rightarrow av

$\theta \rightarrow$ model
parameters

$\alpha \mathbb{N}$

alarche

Vanishing

→ model weight →

→ learning rate ↓↓

→ parameter ↓↓
update S



① Zero

② Random ⑤ No

③ Xavier / Glorot

④ He Uniform

normalized

extended

Random Nonma

Random Unif

1

OTUM

Xavier Uniform



Sigmoid

Regularization

training ↓
validation ↓↓↓ } Over

-n fitting

Normalization

The initialization

ReLU / Leaky ReLU

$\otimes \rightarrow [\text{model}]$ a

n
e Ly

c fixation $\rightarrow \text{O}$

Optimization

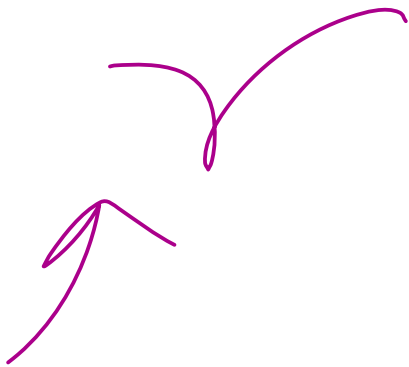
① GD

② SGD

③ Mini-batch

④ Adam ⑥ N

⑤ AdaDelta,



Lumentum

Adegnad

Dropout

Dropout

$$A(0.8)$$