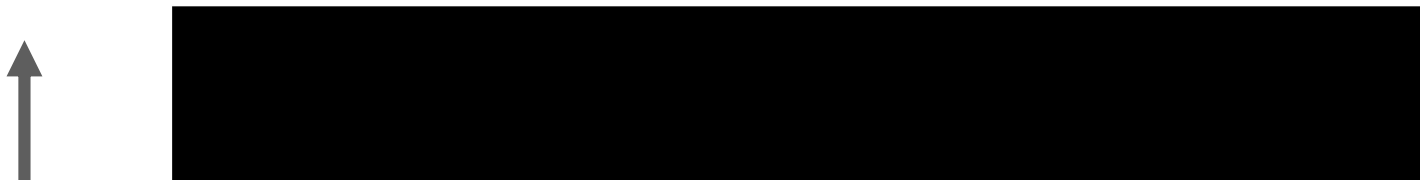
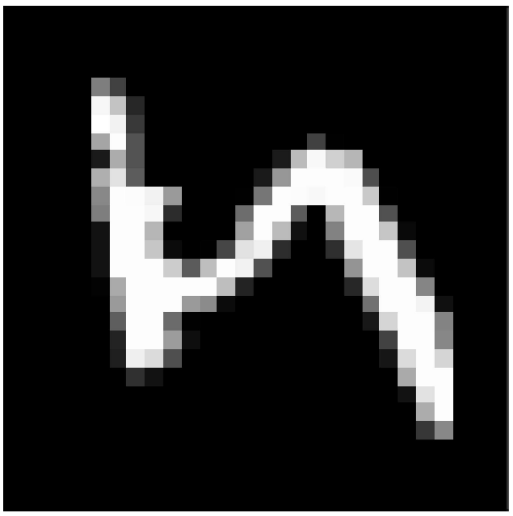
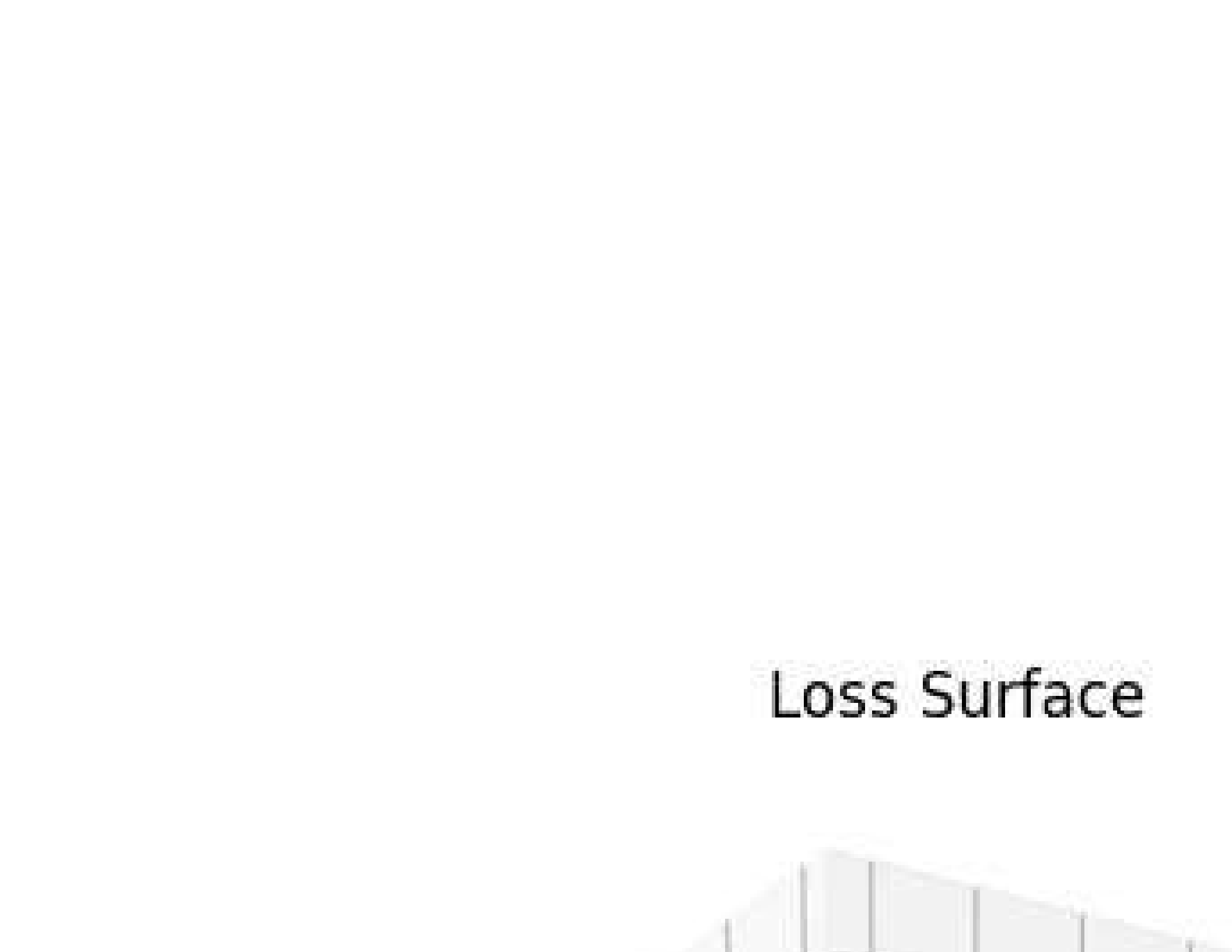
Fundamentals of Deep Learning

Loaded and visualized





A Simpler Model


Loss Surface

6
5

6 5

6 5

The Gradient

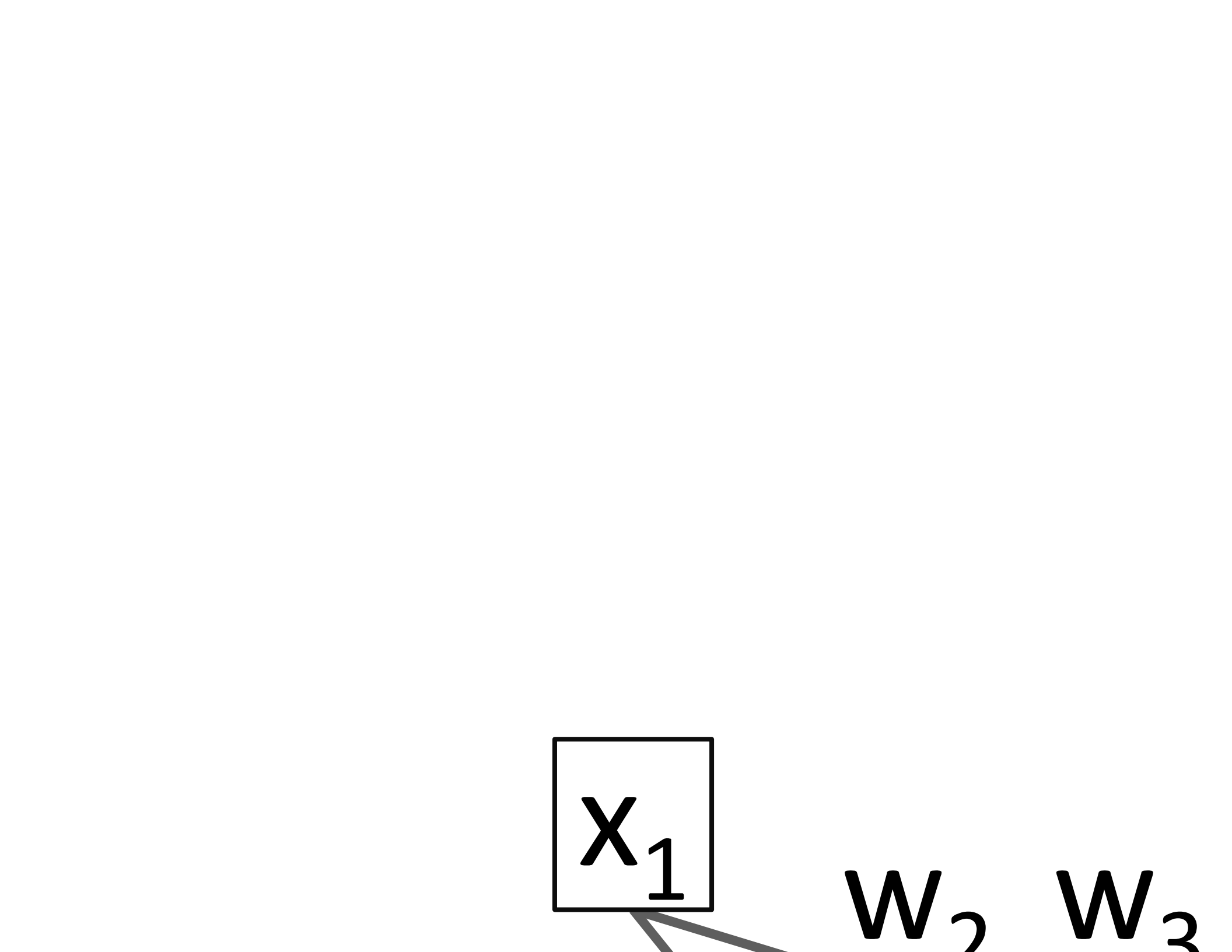
Which direction l
most

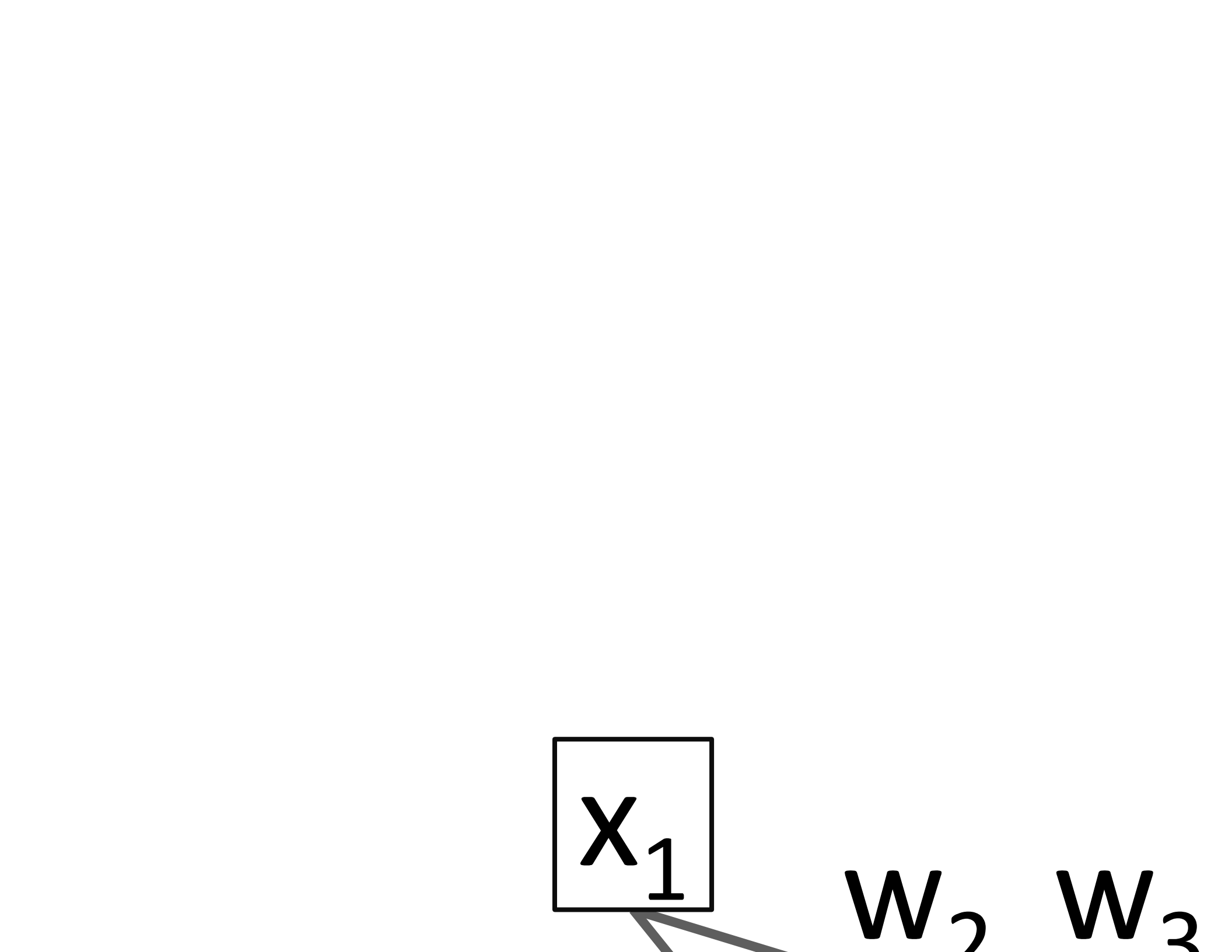
The Gradient

Which direction l
most

Loss – Mo

From Neuron to N


x_1
w_2 w_3

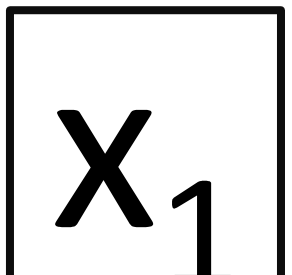

x_1
w_2 w_3

Activation Functio

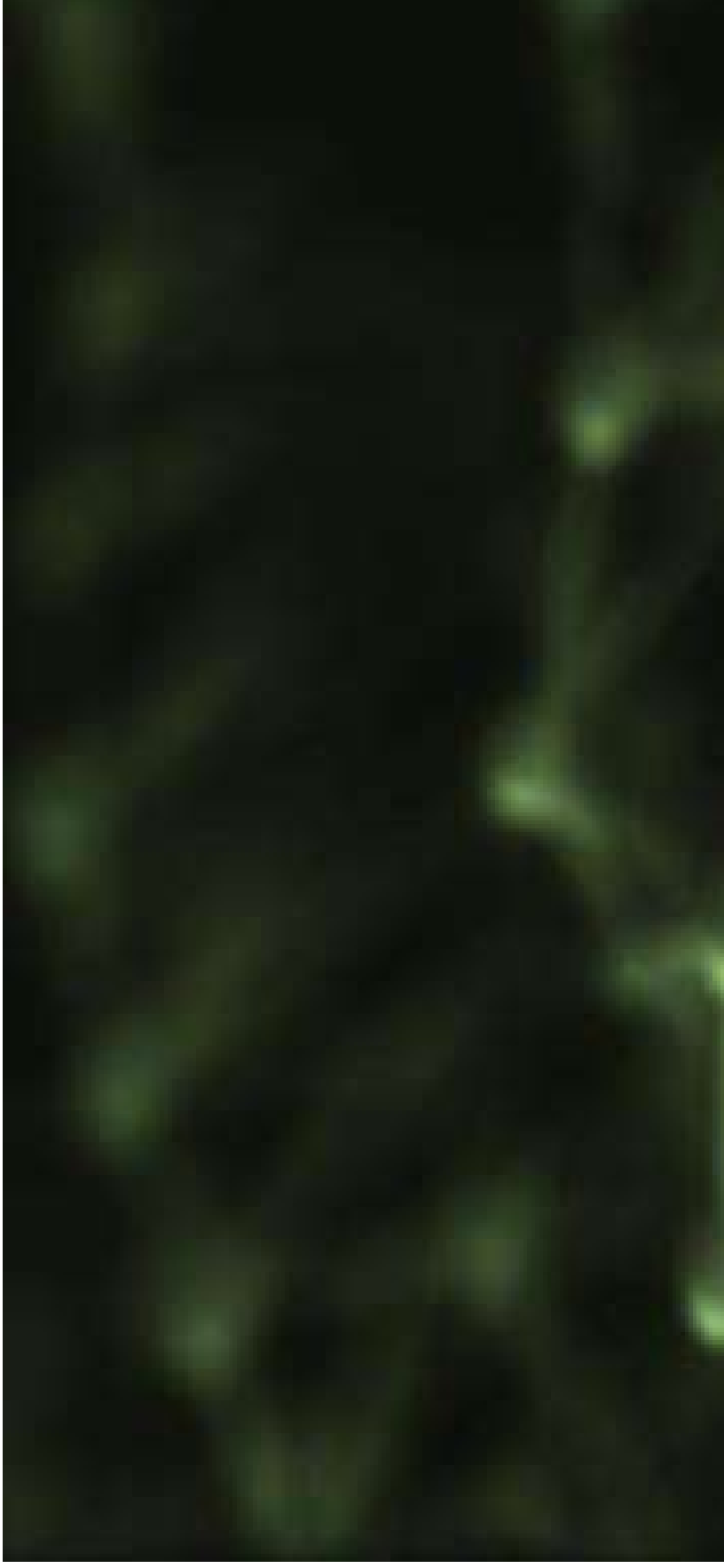
Linear

$$\hat{y} = wx + b$$

Linear


x_1

Overfitting



0.9
0.8

0.8

0.9

Training vs Validatio

Avoid memorization

Training data

From Regression t

4
3

4
3

4
3

4
3

4
3

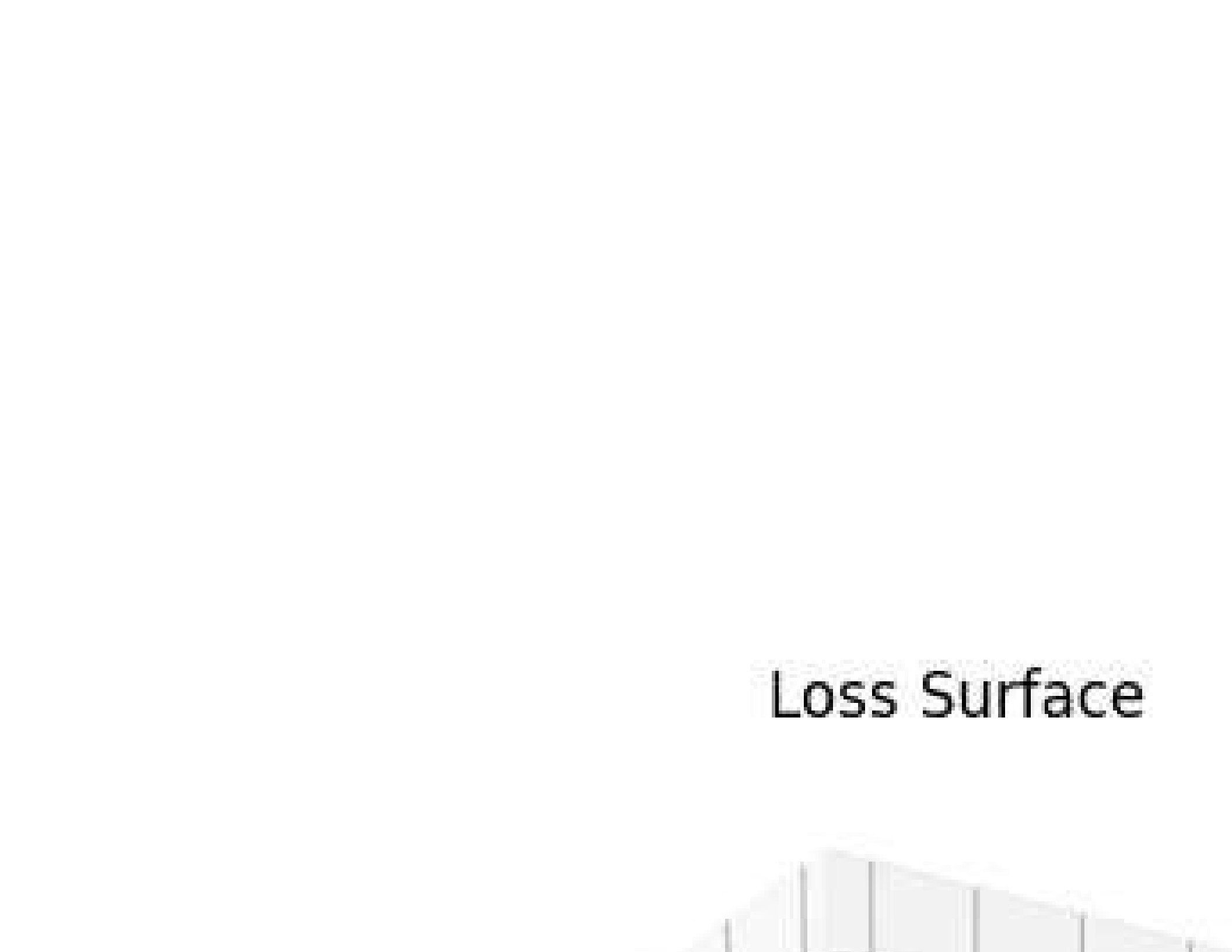
Bringing it Togeth

ก

Let’s go!

Appendix: Grad

$$MSE = \frac{1}{n}\sum_{i=1}^{n}($$


Loss Surface

$$\frac{\partial MSE}{}$$

$$\frac{\partial MSE}{}$$

$$\frac{\partial MSE}{}$$

$$\frac{\partial MSE}{}$$

NVIDIA