Bayesian Neural Networks

Practical Deep Learning for Science
04 July, 2023

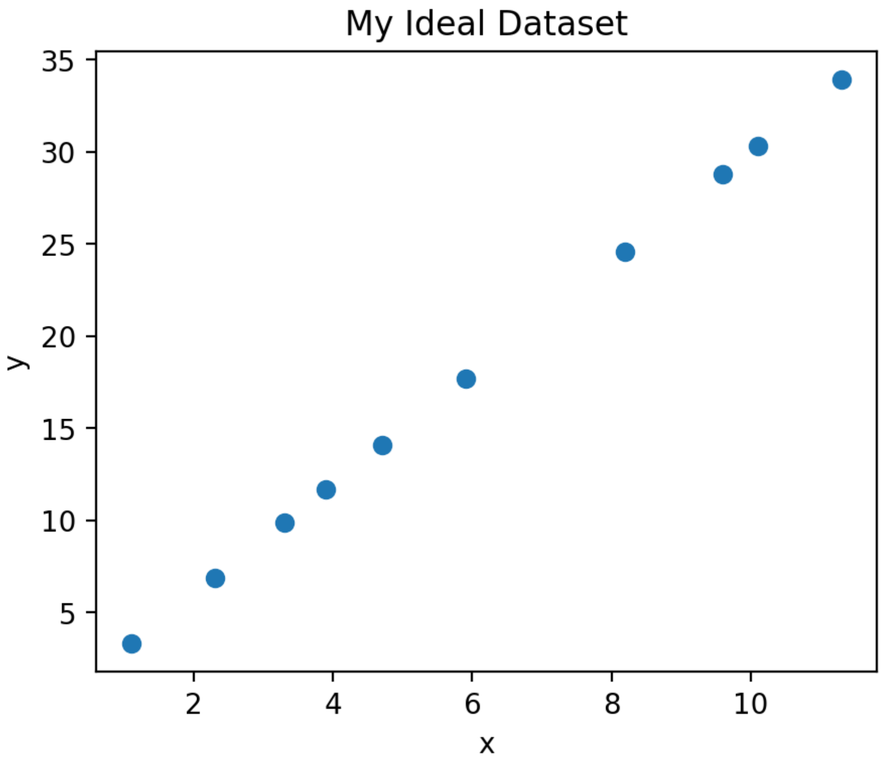
מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

- Nilotpal

Let’s say we are doing an experiment

X	Y
1.1	3.3
2.3	6.9
3.3	9.9
3.9	11.7
9.6	28.8
10.1	30.3
11.3	33.9
4.7	14.1
8.2	24.6
5.9	17.7

What is Y at x = 7?

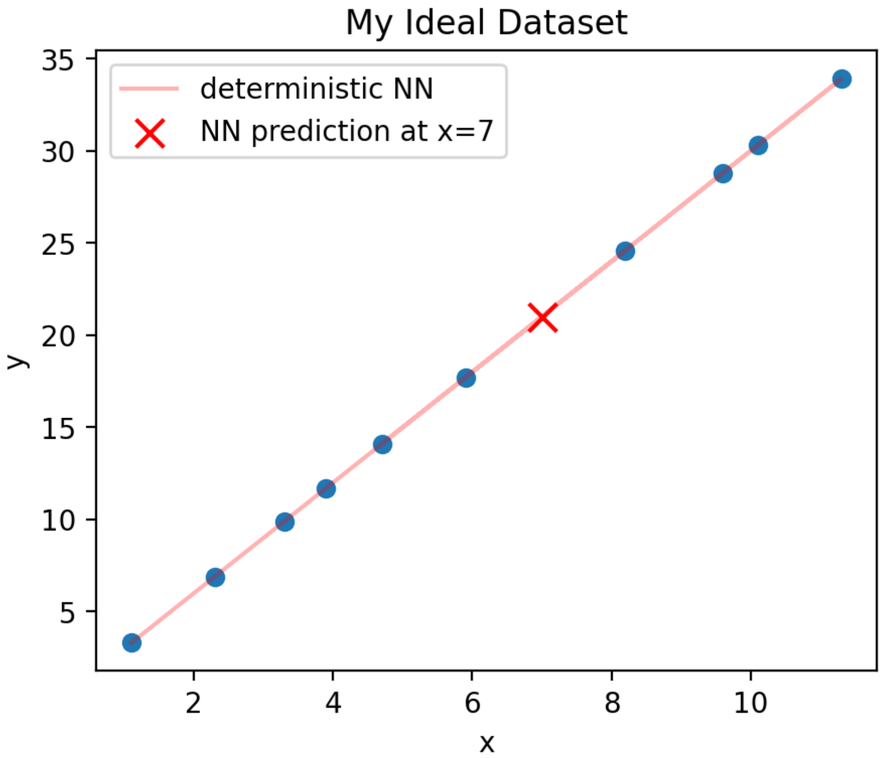


- We can still train an NN with **only one weight** to do the same

Let’s say we are doing an experiment

X	Y
1.1	3.3
2.3	6.9
3.3	9.9
3.9	11.7
9.6	28.8
10.1	30.3
11.3	33.9
4.7	14.1
8.2	24.6
5.9	17.7

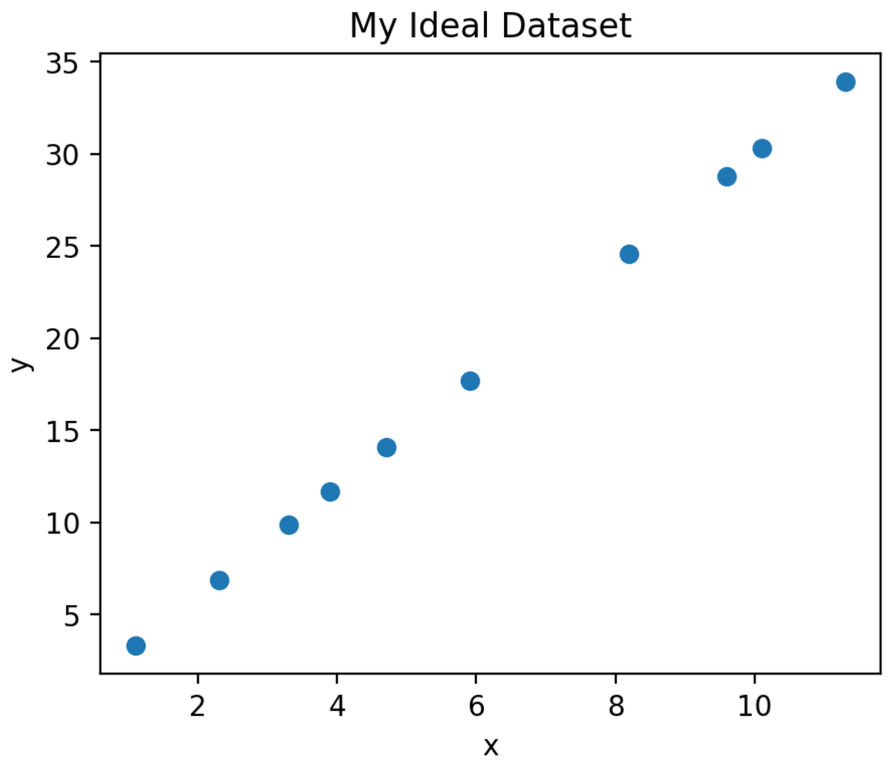
What is Y at x = 7?



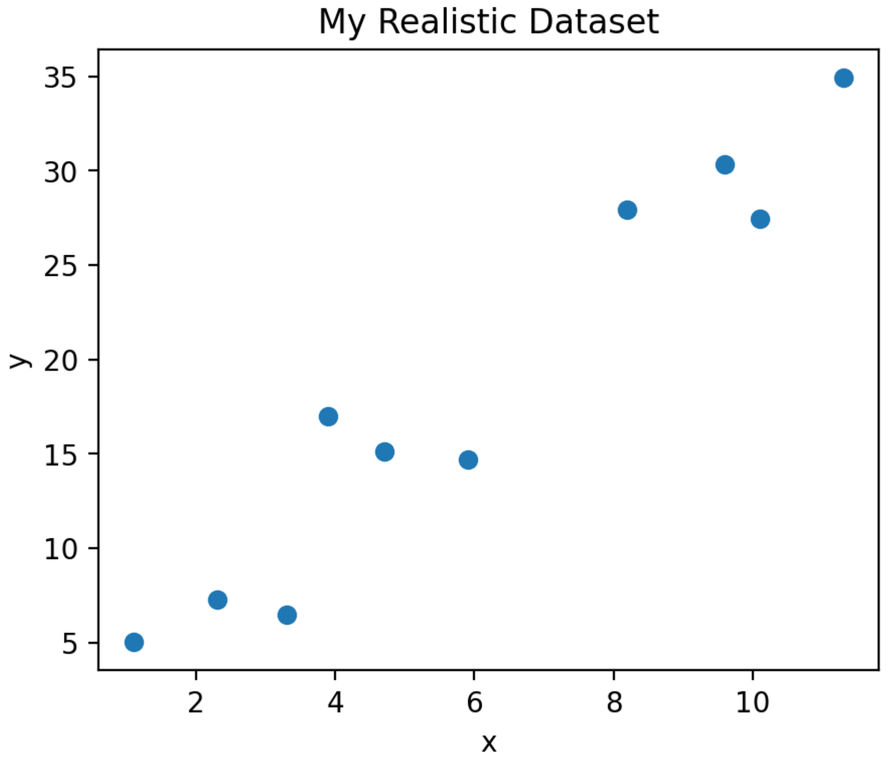
- We can still train an NN with **only one weight** to do the same

Let’s say we are doing an experiment

Ideal situation

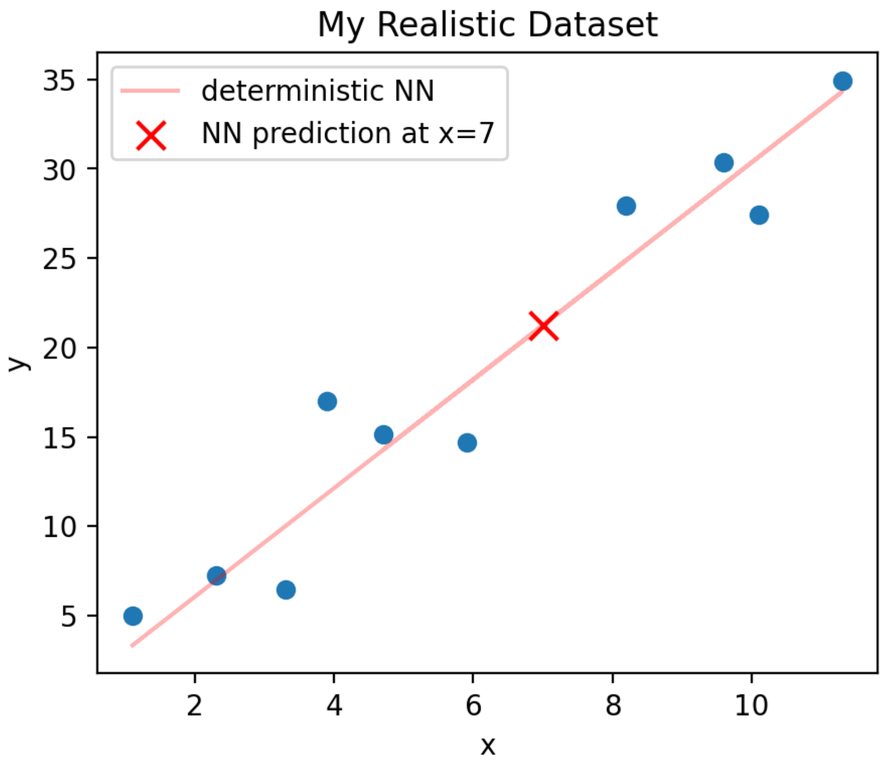


What we see

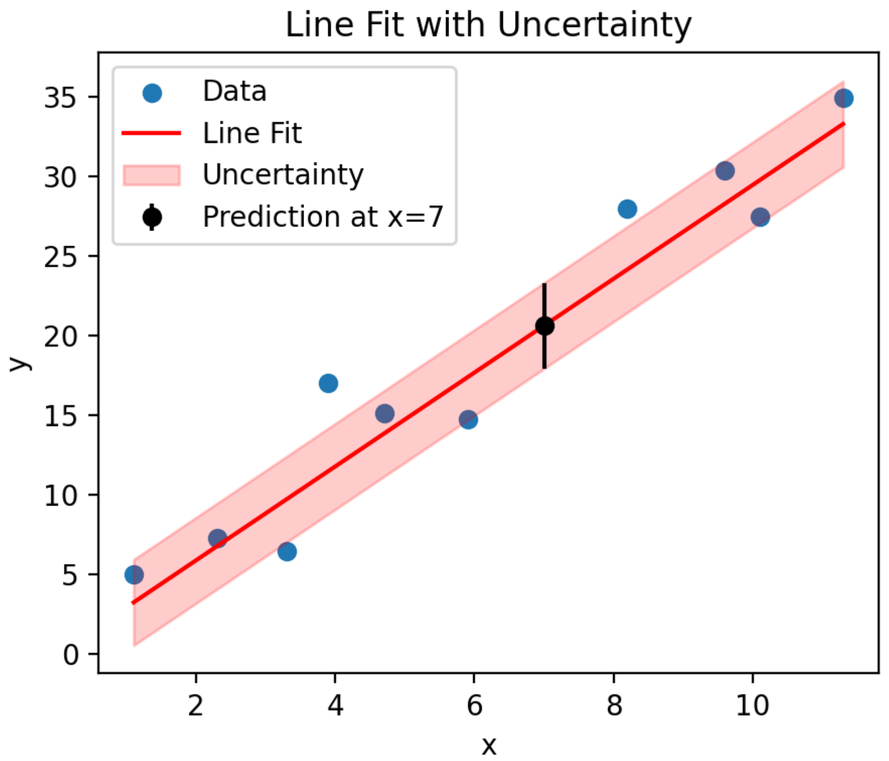


If we train a network

We will get



What we want



How do we quantify the uncertainty with NN

Aleatoric uncertainty

- Noise in the data
- More data won't help

vs

Epistemic uncertainty

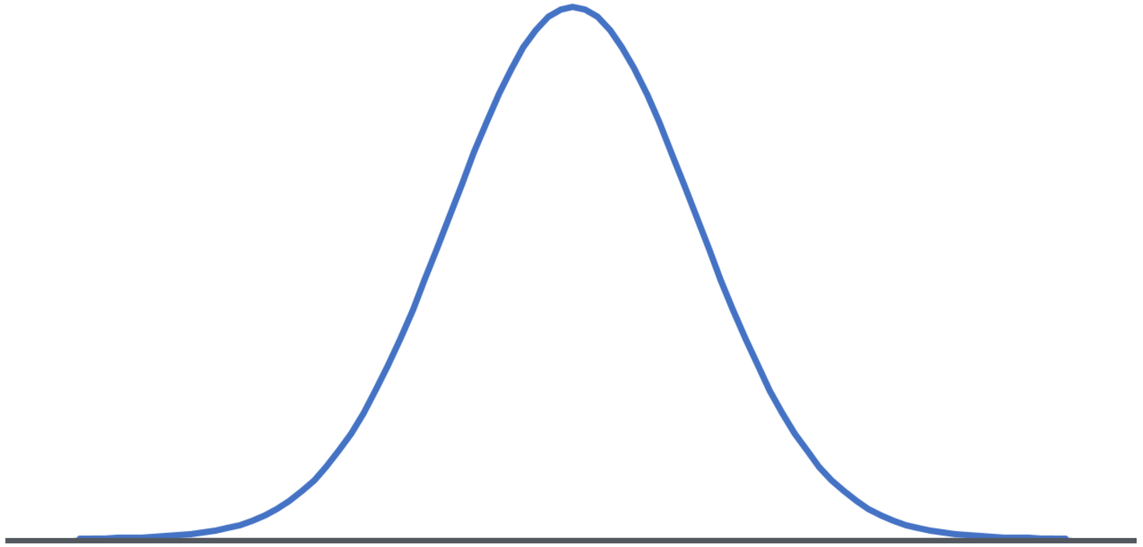
- Uncertainty in the model parameters
- More data will help

- We need to modify the network to accommodate these two uncertainties

Model uncertainty

Model weight before

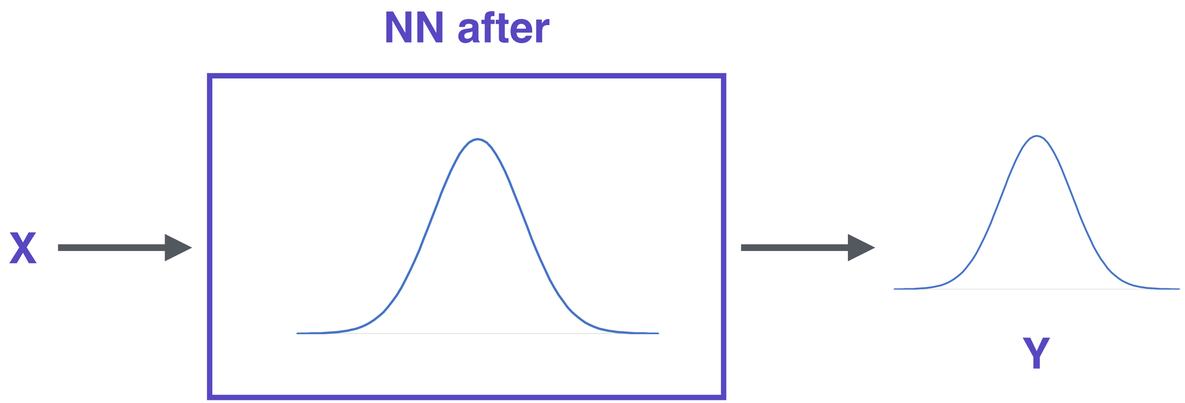
Gaussian (μ, σ)



Model weight after

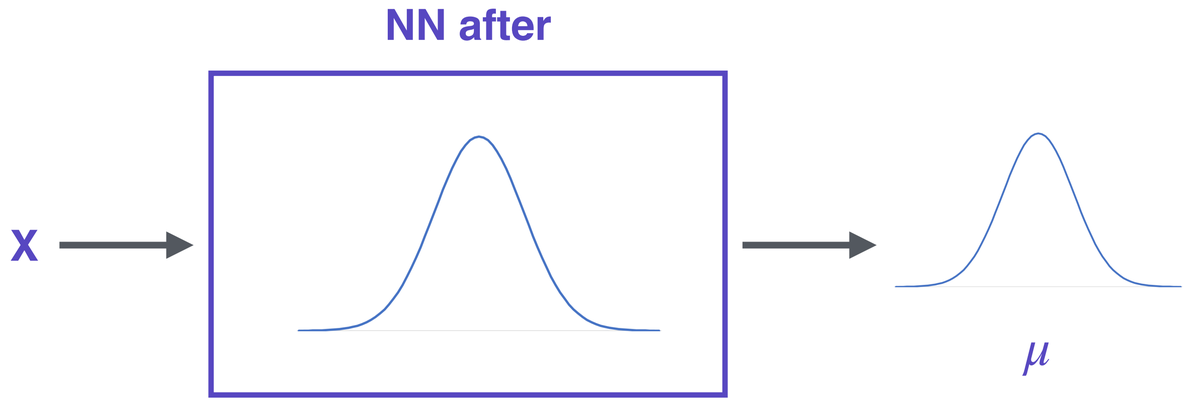
- Each weight, instead of being a number, will be two numbers - μ and σ
- During forward pass, we sample from the Gaussian (μ, σ)

Model uncertainty



How about the uncertainty from the data

- Let's add another Gaussian into the picture
- Mean = output of the NN (a gaussian)
- Sigma = another Gaussian we'll learn



$\hat{Y} =$ **Gaussian(mu, sigma)**