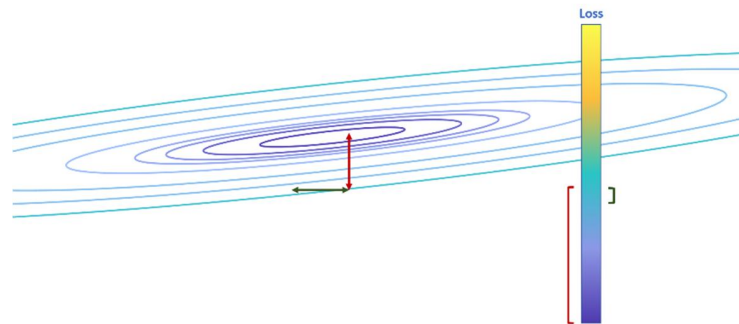# Training Options

A basic understanding of how training works can help you sift through the many training options available. The next few pages will break down some of the important ones:

- Mini-batches
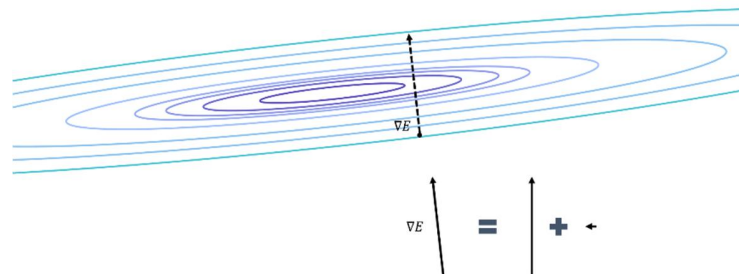- Learning rates
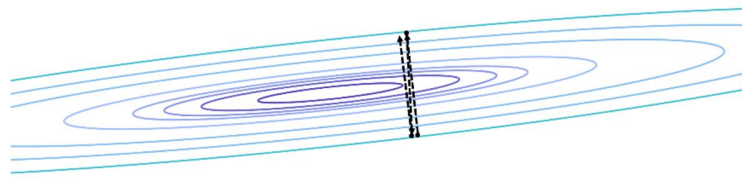- Gradient clipping
- Algorithms

Gradient Clipping

1.

A common problem with using gradients is that the loss function may change rapidly with some parameters but slowly with others.
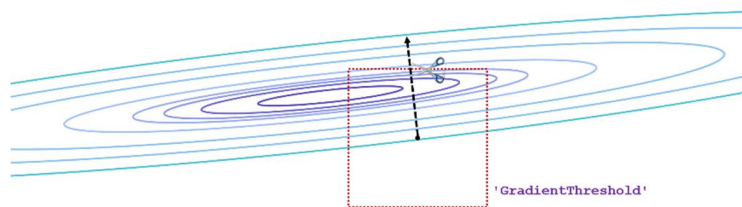
2.

This means some values of the gradient are large and some are small.

3.

Very large gradient values can make your parameter values change wildly.

4.

'GradientThreshold'

A simple way to control this is to put a maximum size on the gradient values.