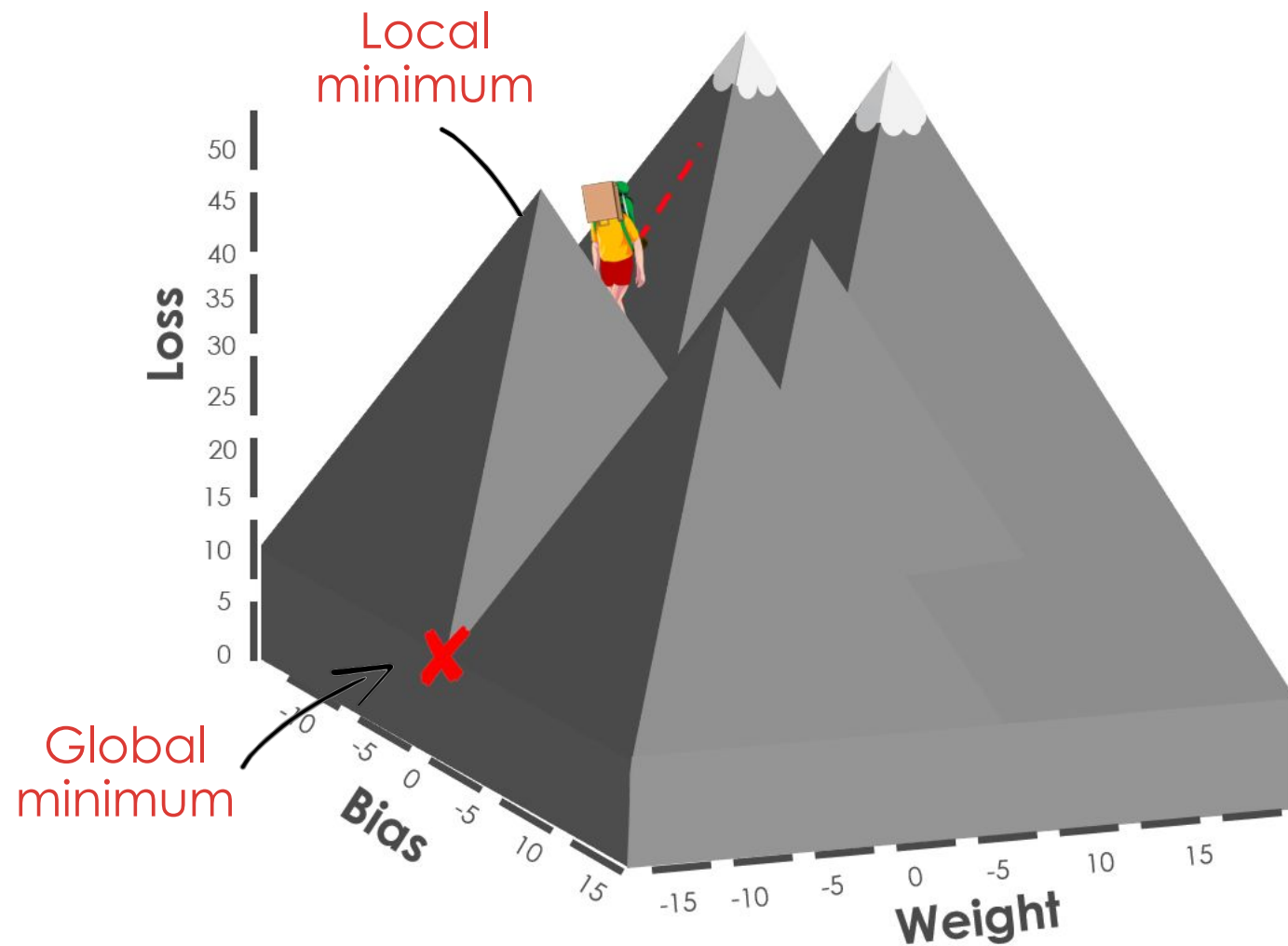




Gradient Descent Optimization

Deep Learning Pre-Work

Local Minimums



This file is meant for personal use by muratgguzel77@yahoo.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Agenda

- Epochs
- Batch size
- Learning rate
- Optimizers

Epochs

This file is meant for personal use by muratgguzel77@yahoo.com only.

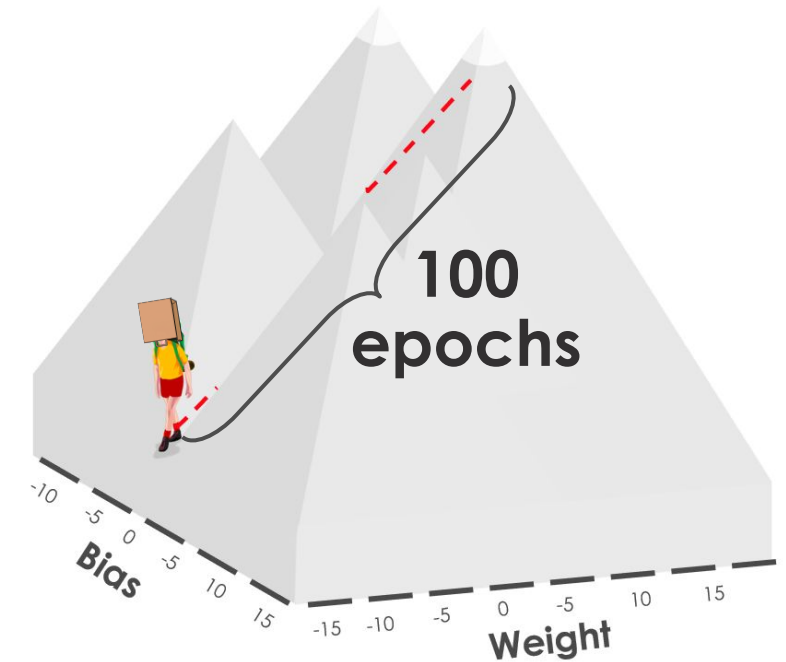
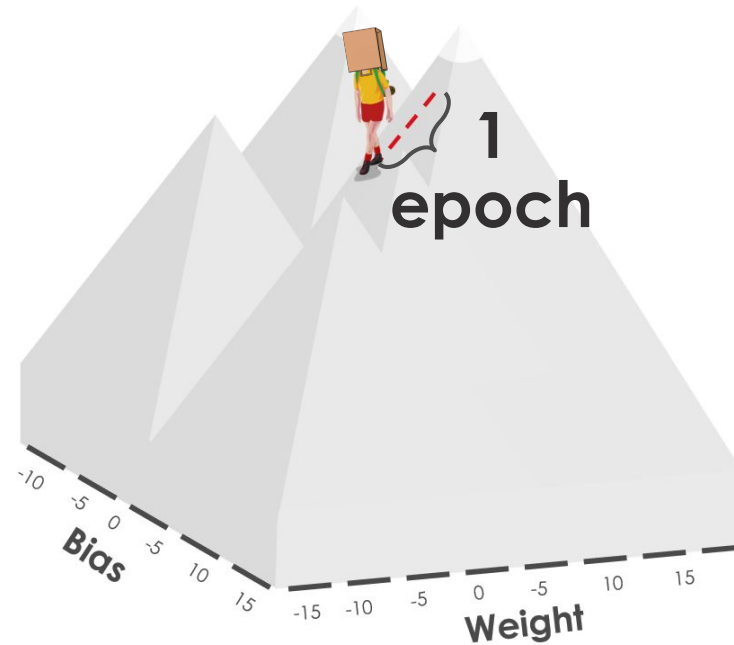
Sharing or publishing the contents in part or full is liable for legal action.

Epochs

Training Data

0	0.0	0.0	1.00000
1	1.0	0.0	1.00000
2	1.0	0.0	0.67065
3	1.0	0.0	0.97467
4	1.0	0.0	0.28409
...
240	0.0	0.0	1.00000
241	1.0	0.0	0.00000
242	1.0	0.0	1.00000
243	1.0	0.0	0.94701
244	0.0	0.0	0.00000

245 rows × 35 columns



This file is meant for personal use by muratgguzel77@yahoo.com only.

Sharing or publishing the contents in part or full is liable for legal action.

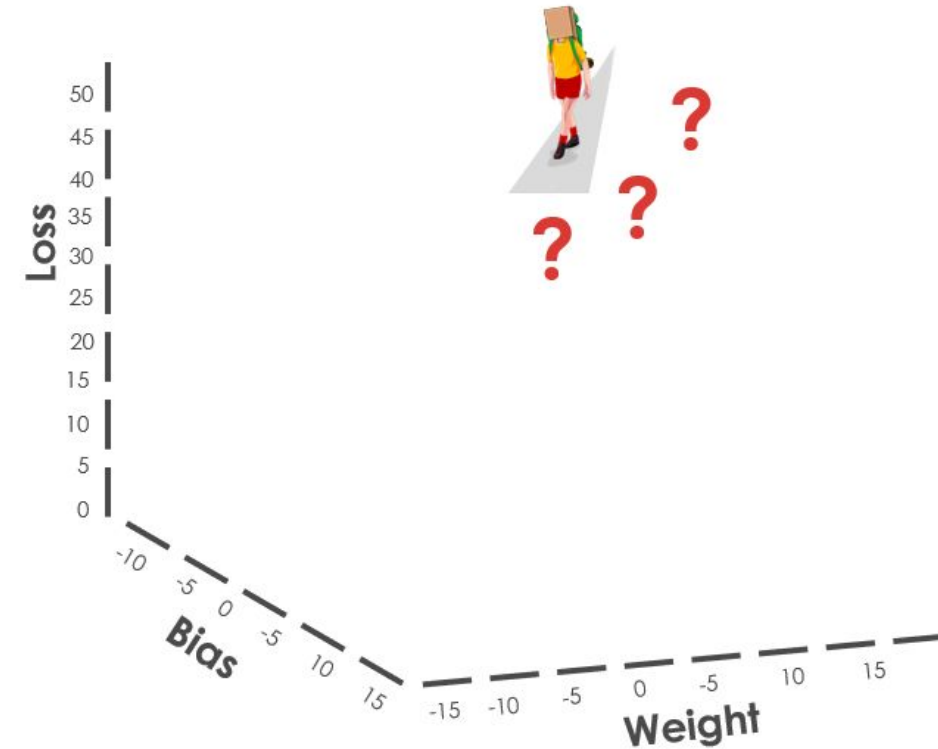
Batch Size

This file is meant for personal use by muratgguzel77@yahoo.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Gradient Descent Steps

- Step 1 Start at a random bias and weight and calculate the loss
- Step 2 Take a step in the direction with the steepest gradient**
- Step 3 Calculate the new loss
- Step 4 Repeat steps 2 and 3



Batch size

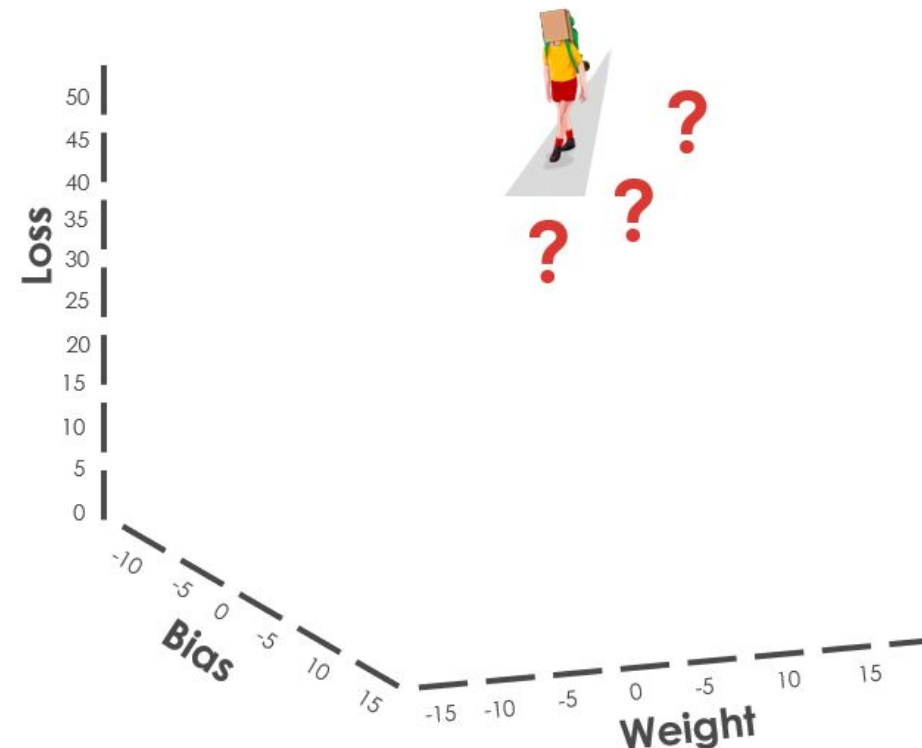
**Batch
Gradient Descent**

**Mini-Batch
Gradient Descent**

**Stochastic
Gradient Descent**

0	0.0	0.0	1.00000
1	1.0	0.0	1.00000
2	1.0	0.0	0.67065
3	1.0	0.0	0.97467
4	1.0	0.0	0.28409
...
240	0.0	0.0	1.00000
241	1.0	0.0	0.00000
242	1.0	0.0	1.00000
243	1.0	0.0	0.94701
244	0.0	0.0	0.00000

245 rows × 35 columns



This file is meant for personal use by muratgguzel77@yahoo.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Batch size

Batch Gradient Descent

Batch size: 245



Mini-Batch Gradient Descent

Batch size: 35



Stochastic Gradient Descent

Batch size: 1



Training sample size: 245
Epochs: 10

This file is meant for personal use by muratgguzel77@yahoo.com only.

Sharing or publishing the contents in part or full is liable for legal action.

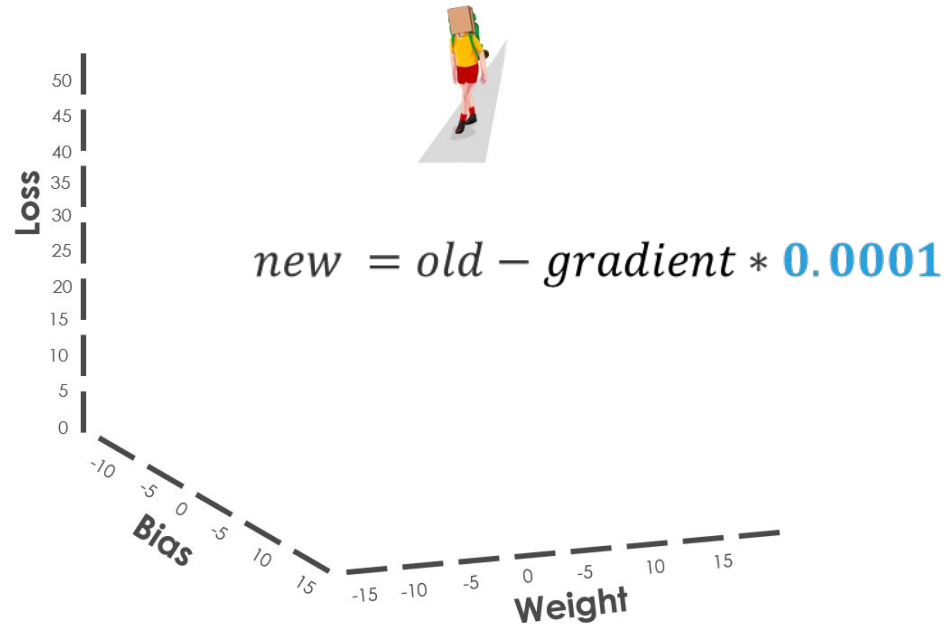
Learning Rate

This file is meant for personal use by muratgguzel77@yahoo.com only.

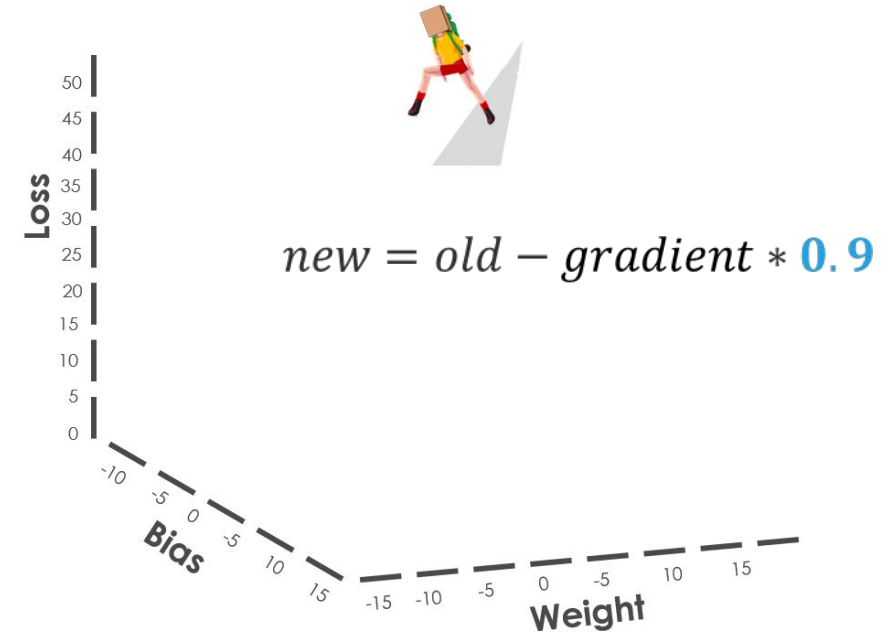
Sharing or publishing the contents in part or full is liable for legal action.

Learning Rate

Little step



Big step



This file is meant for personal use by muratgguzel77@yahoo.com only.

Sharing or publishing the contents in part or full is liable for legal action.

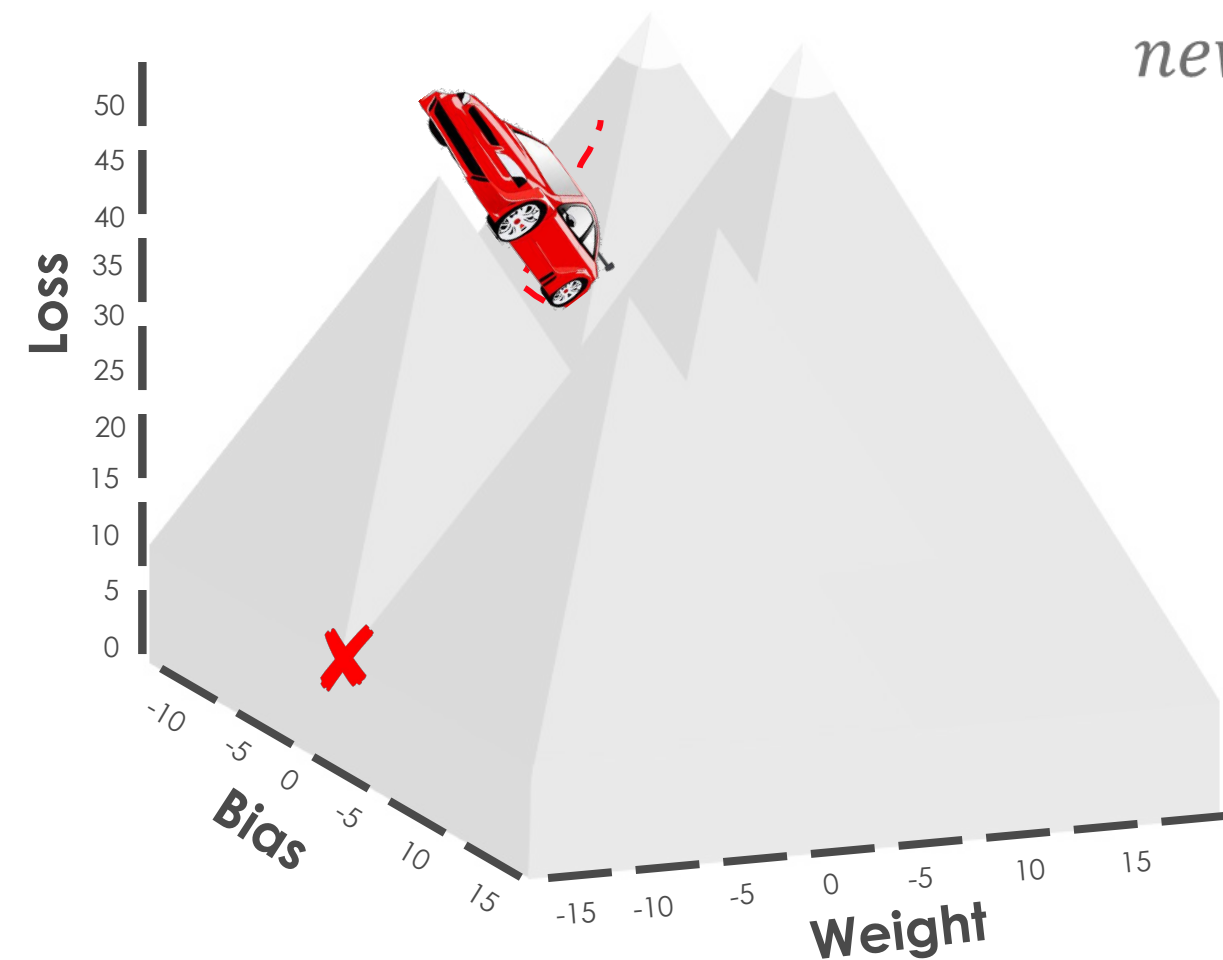
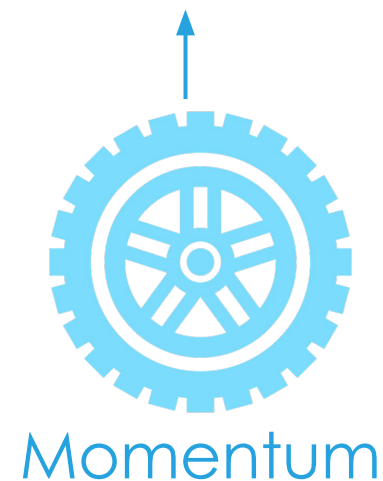
Optimizers

This file is meant for personal use by muratgguzel77@yahoo.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Momentum

$$\text{new} = \text{old} - \text{gradient} * \text{learning rate}$$

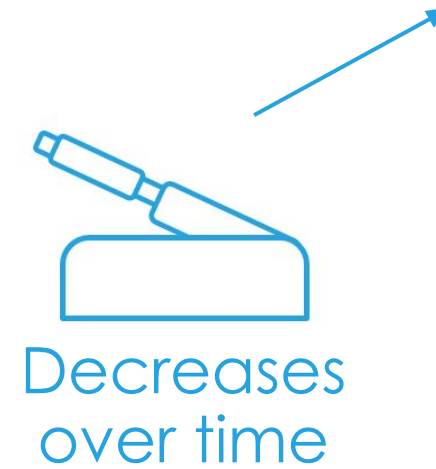
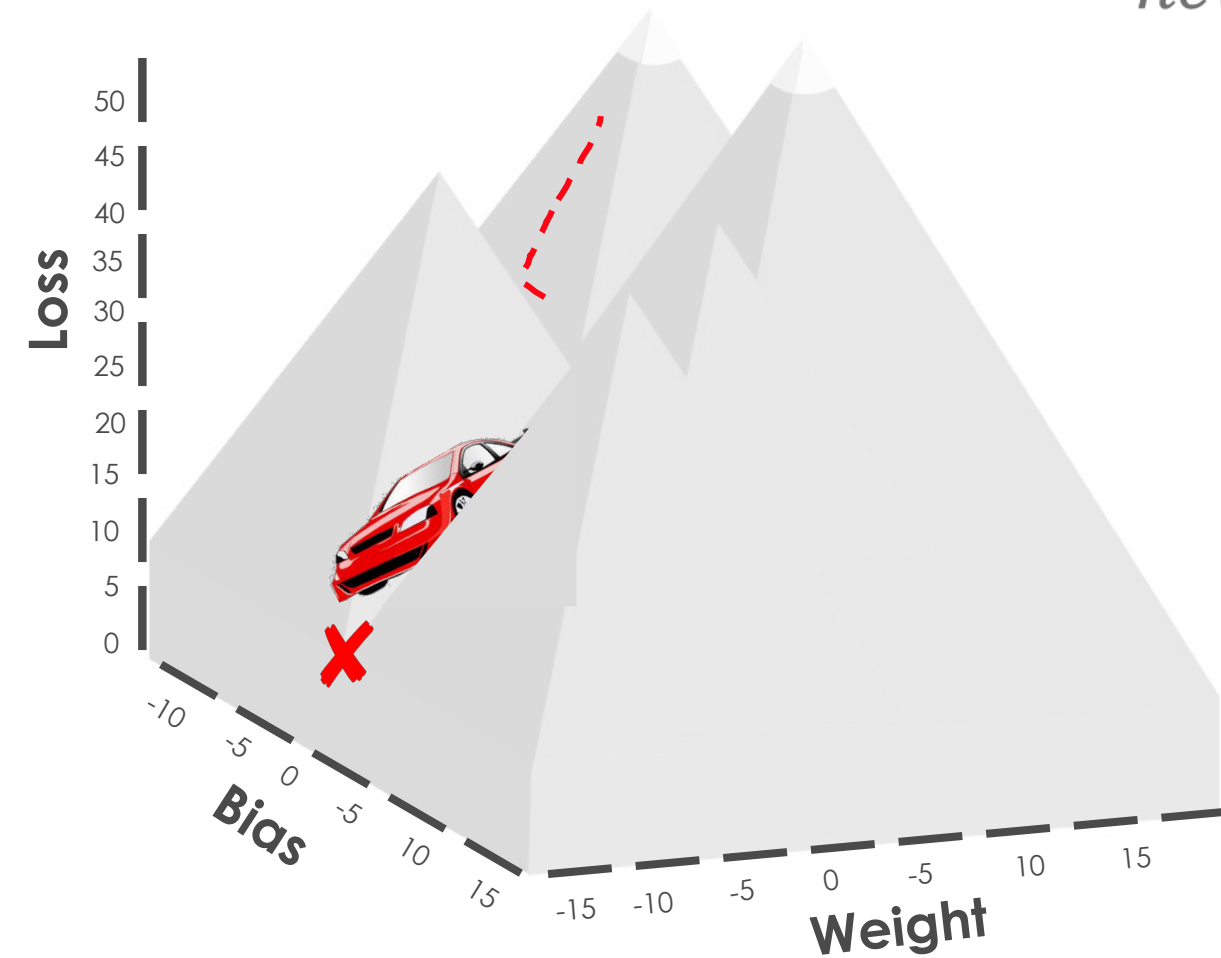


This file is meant for personal use by muratgguzel77@yahoo.com only.

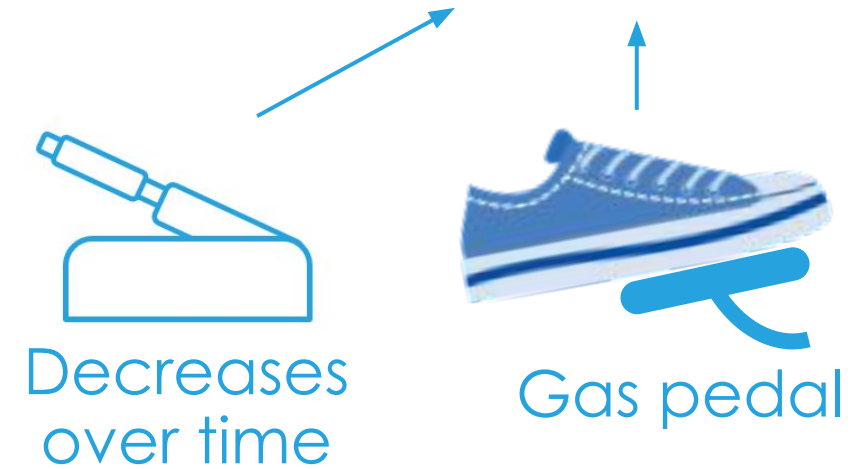
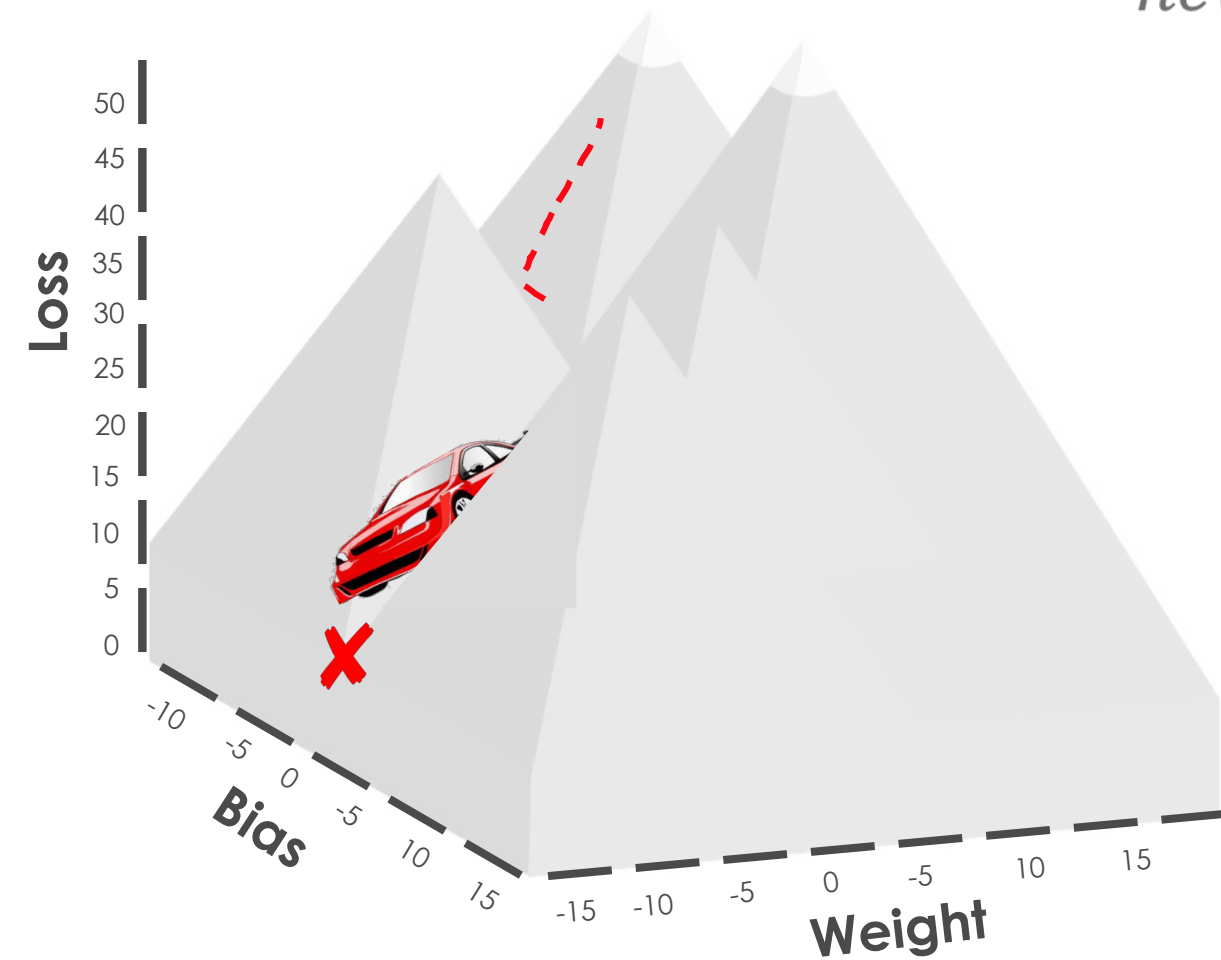
Sharing or publishing the contents in part or full is liable for legal action.

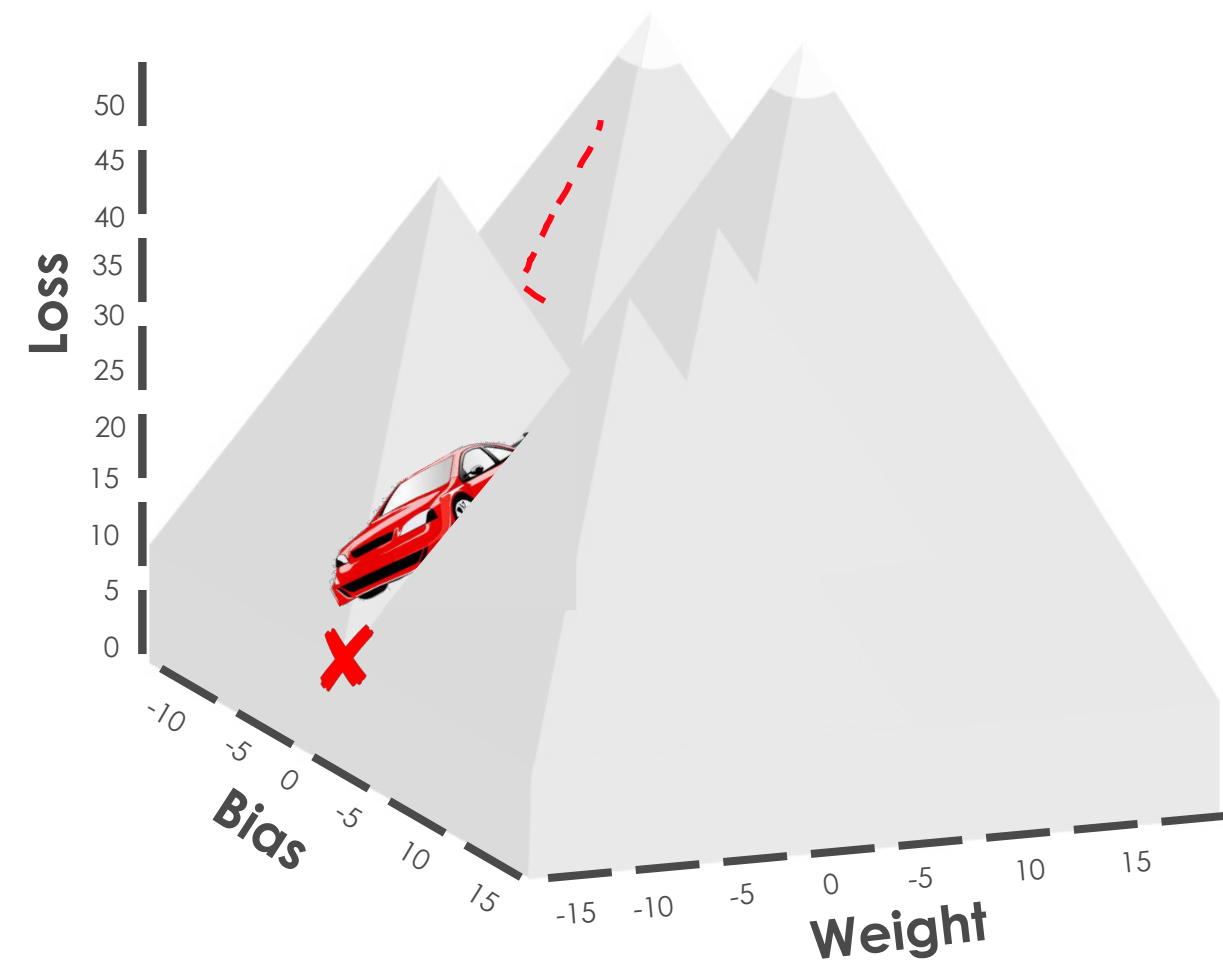
Adagrad

$$\text{new} = \text{old} - \text{gradient} * \text{learning rate}$$

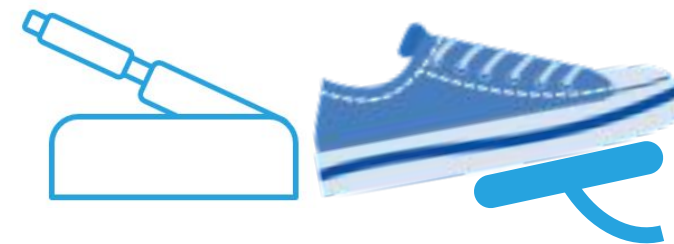


$$new = old - gradient * learning\ rate$$





Momentum



RMSprop