

- **Conclusion of Nesterov - Acceleration - Gradient , Momentum – Gradient and Vanilla-Gradient descents:**

Weight(w)	Bias(b)	Learning Rate(eta)	Epochs(epoch)	Vanilla-Gradient-Error	Momentum-Gradient-Error	Nesterov-Gradient-Error	Vanilla-Gradient-Time	Momentum-Gradient-Time	Nesterov-Gradient-Error
-2	-2	5	700	1.0e <sup>-31</sup>	3.73e <sup>-32</sup>	3.73e <sup>-32</sup>	0.015625	0.03125	-0.03125
-2	-2	10	700	2.5e <sup>-32</sup>	7.7e <sup>-33</sup>	6.54e <sup>-33</sup>	0.03125	0.015625	-0.015625
-2	-2	15	700	7.7e <sup>-33</sup>	9.6e <sup>-33</sup>	7.7e <sup>-33</sup>	0.03125	0.03125	-0.03125

- Here, I have taken learning rate as a variable and check for the same weight, Bias, Epochs.
- The output here is, that the nesterov gradient is faster than the other gradients.
- Time of nesterov is low other than the two gradients and error rate of nesterov is less than the other gradient descents.
- As in the last row, the error rate of the vanilla and nesterov gradient is similar and the error rate of the momentum gradient is comparatively higher. But the time of the nesterov is lower among all three gradients.
- So we can say that the Nesterov-Acceleration Gradient is faster among all the three gradients.