|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **W** | **B** | **Learning rate** | **Epochs** | **Error** |
| -2 | -2 | 1.0 | 100 | 0.017741009621761353 |
| -2 | -2 | 1.0 | 1000 | 2.4291865713546816e-11 |
| -2 | -2 | 1.0 | 2000 | 2.3947265822163888e-20 |
| -2 | -2 | 3.0 | 100 | 5.716362296559683e-05 |
| -2 | -2 | 5.0 | 100 | 7.114464162870803e-07 |
| -2 | -2 | 7.0 | 100 | 6.461558886290009e-09 |
| -2 | -2 | 10.0 | 100 | 1.844555021169613e-12 |
| -2 | -2 | 12.0 | 100 | 1.3090084531435982e-14 |
| -2 | -2 | 13.0 | 100 | 3.005477552584132e-13 |
| -2 | -2 | 14.0 | 100 | 1.64805147430089e-17 |
| 0 | 0 | 1.0 | 100 | 0.0013755069444627295 |
| 5 | 5 | 1.0 | 100 | 0.32453212800819825 |
| 10 | 10 | 1.0 | 100 | 0.3249997552707788 |
| 1.8 | 0 | 1.5 | 1000 | 1.0305100703007212e-16 |

**Conclusion:**

* From the first 3 row, we can see that as we increase the epochs, the error is decreased.
* Also the weights are not becoming constant for this steps i.e. we get same decimals up to 6th of 7th position after that they vary.
* From row 4 to 10 we can see that if we change only learning rate then it affect the error too. As we increased learning rate from 1.0 to 12.0 and keeping weight, bias and epochs’ constant then error is decreased but if we change learning rate to 13.0 then error is increased and after that we change learning rate to 14.0 then error is still decreasing. So we can say that if we give learning rate too large then there might be overshoot and error become large.
* From last 4 row we can see that different weight and bias value at the time of initialize affect the error too. If we initialize weight and bias to very large value then it will take more epochs to converge.
* It is depend upon the problem that how to initialize weight, bias, and what value for the learning rate and epochs. If we give too much value to the learning rate then it might overshoot and if we give very small value to learning rate then it will require more epochs to converge and it takes more time to train.

Now for data set X = [0.5, 1.25, 2, 2.5, 3.75, 4.2] and Y = [0.2, 0.45, 0.65, 0.8, 0.95, 1]

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **W** | **B** | **Learning rate** | **Epochs** | **Error** |
| -2 | -2 | 1.0 | 100 | 0.00128719039079461 |
| -2 | -2 | 1.0 | 1000 | 0.0009055798154162559 |
| -2 | -2 | 1.0 | 2000 | 0.0009055798154162556 |
| -2 | -2 | 3.0 | 100 | 0.0009056217184718256 |
| -2 | -2 | 5.0 | 100 | 0.03311773499453983 |
| -2 | -2 | 7.0 | 100 | 0.037976472275588924 |
| -2 | -2 | 10.0 | 100 | 0.11320812426842726 |
| -2 | -2 | 12.0 | 100 | 0.056744685552190116 |
| -2 | -2 | 13.0 | 100 | 0.5370548042491412 |
| -2 | -2 | 14.0 | 100 | 0.049705056730003014 |
| 0 | 0 | 1.0 | 100 | 0.0010382414239020284 |
| 5 | 5 | 1.0 | 100 | 0.5532736989476383 |
| 10 | 10 | 1.0 | 100 | 0.5537497551776843 |
| 1.8 | 0 | 1.5 | 1000 | 0.0009055798154162593 |

**Conclusion:**

* From the first three rows we can derive that for small dataset error gets very very small and weight and bias don’t get constant while increasing number of epochs and keeping everything else same whereas, for bigger data set as only the number of epochs increases the weights gets constant but the error decreases very slowly.
* From row 4 to 10 we can see that if we change only learning rate then it affect the error too. As we increased learning rate from 1.0 to 10.0 and keeping weight, bias and epochs’ constant then error is increased but if we change learning rate to 12.0 then error is decreased and after that we change learning rate to 13.0 then error is increasing and again after changing it to 14.0 it decreases again. So we can say that if we give learning rate too large then there is overshoot and error become large.
* From last 4 row we can see that different weight and bias value at the time of initialize affect the error too. If we initialize weight and bias to very large value then it will take more epochs to converge.