ITE4053 Deep Learning course

Practice 4

2016026080 Dajin Han

* **Results**

|  |  |  |
| --- | --- | --- |
|  | BinaryCrossentropy | MeanSquaredError |
| Acc\_train | 0.984600 | 0.985200 |
| Acc\_test | 0.987000 | 0.982000 |

|  |  |  |  |
| --- | --- | --- | --- |
|  | SGD | RMSprop | Adam |
| Acc\_train | 0.983900 | 0.985300 | 0.940100 |
| Acc\_test | 0.988000 | 0.973000 | 0.931000 |
| Time\_train | 5.122176 | 6.413932 | 5.869797 |
| Time\_test | 0.249455 | 0.257780 | 0.249931 |
| Loss type | Binary CE | Binary CE | Binary CE |

|  |  |  |  |
| --- | --- | --- | --- |
|  | Python Result (local) | CPU version (local) | GPU version (server) |
| Acc\_train | 0.984600 | 0.984700 | 0.985400 |
| Acc\_test | 9.770000 | 0.986000 | 0.987000 |
| Time\_train | 0.8349459 | 4.460400 | 5.611204 |
| Time\_test | 0.0001789 | 0.341082 | 0.180154 |
| Loss type | CE | CE | CE |
| Optimizer type | SGD | SGD | SGD |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Mini-batch=1 | Mini-batch=32 | Mini-batch=128 | Mini-batch=1000 |
| Acc\_train |  |  |  |  |
| Acc\_test |  |  |  |  |
| Time\_train |  |  |  |  |
| Time\_test |  |  |  |  |
| Loss type |  |  |  |  |
| Optimizer type |  |  |  |  |

* **What I learned**

Generally, Adam optimizer makes best performance. But explain it is because of the activation function sigmoid. Initial distribution of weight value can cause the local minimum. Maybe I need to study more about this.