**优化器对比**

（lr=0.1 epoch=500 batch=32）

表1 loss图对比

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sgd | Sgdm | Adagrad | Rmsprop（后面epoch的loss都是nan，建议调小学习率） | Adam |
|  |  |  |  |  |

表2 ACC图对比

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sgd | Sgdm | Adagrad | Rmsprop | Adam |
|  |  |  |  |  |

表3 训练耗时（total\_time）对比

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sgd | Sgdm | Adagrad | Rmsprop | Adam |
| 12.62293004989624 | 13.419726133346558 | 13.302583694458008 | 14.12432050704956 | 19.189162492752075 |

**优化器对比**

（lr=0.01 epoch=100 batch=32）

表1 loss图对比

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sgd | Sgdm | Adagrad | Rmsprop | Adam |
|  |  |  |  |  |

表2 ACC图对比

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sgd | Sgdm | Adagrad | Rmsprop | Adam |
|  |  |  |  |  |

表3 训练耗时（total\_time）对比

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sgd | Sgdm | Adagrad | Rmsprop | Adam |
|  |  |  |  |  |