**优化器对比**

（lr=0.1 epoch=500 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 |
| 2.9791901111602783 | 3.2608582973480225 | 3.170058250427246 | 3.277435064315796 | 3.7047104835510254 |
|  |  |  |  |  |
| 运行时间最短，复杂度最低  但是该参数 调参较为困难 |  |  |  | 复杂度最高  但是基本是没有什么调参 |

**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 |
|  |  |  |  |  |