本次数据挖掘比赛总结

本次使用模型尝试及效果:

1. model:baseline(3层FC) loss(经典loss)(mean l2)

效果56.4 0.9118(分别最优的结果)

上传结果 39.8(历代最优的rank) 0.959

作用:长期作为模型融合的低rank部分

1. model:baseline(3层FC) loss(自己的针对惩罚loss )(l2 + 单独惩罚)

效果 67.6 0.978

上传结果 50 0.991(历代最优ontime)

作用:初期作为高ontime的部分

后期作为调整惩罚系数的主流loss

1. model:pro\_baseline(9层FC),loss(自己的针对惩罚loss)

效果: 64 0.977(调整loss惩罚系数之后)

上传结果:46.6 0.982

作用:后期作为模型融合的辅助文件3,也是直到最后最有希望的模型

缺点:由于层数的加深,会产生梯度爆炸和梯度消失的问题

1. model:pro\_base(9层),loss(自己的针对惩罚 smoothl2)

效果:56.40 0.811(无惩罚系数),64.40,0.981(由惩罚系数)

上传结果:46.6 0.982 40.04 0.953

作用:作为后期高ontime的候补之一

缺点:能解决梯度消失,但是无法解决梯度爆炸

1. 加入卷积网络(3层卷积网络, 9层FC),针对惩罚loss

效果:73.2 0.991

经过loss惩罚系数降低后,使用新的loss计算方法(log(2+exp(abs(x[i]-y[i]))))

效果62.5 0.966

长传结果:44.2 0.973

1. 加入卷积的残差网络:

效果:93.5 0.87

…

1. 加入残差的FC网络

效果:66.6 0.966

…

1. Day和hour分开训练:

效果:70.4 0.918

…

9 使用走件时间的双层网络训练:

效果:67.6 0.87

…

另外的:

数据增广的效果不明显

卷积的效果不好

残差网络的拟合速度快,训练速度慢,最终效果不好

每次训练中,会有特定的一部分偏差很大,怀疑是数据有问题,但没有人做数据的统计.