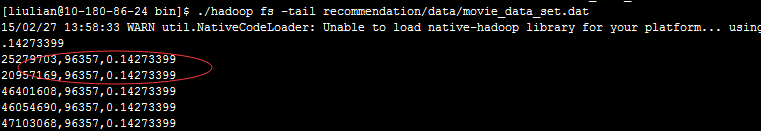
命令使用：

spark-submit --class "data.handler.RatingIDMaping" --driver-memory 10G RatingIDMaping.jar recommendation/data/movie\_data\_set.dat recommendation/data/train\_test 90

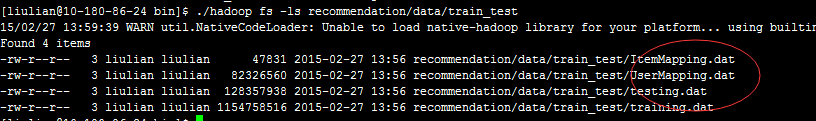
说明：

1 recommendation/data/movie\_data\_set.dat是原始用户对物品的打分

数据格式是



2 recommendation/data/train\_test是结果存放的文件夹



ItemMapping.dat是物品ID的转换前后映射关系

UserMapping.dat是用户ID的转换前后映射关系

testing.dat是用户和物品ID转换后的打分测试集

training.dat是用户和物品ID转换后的打分训练集

3 90是训练集的比例，表示训练集的记录数占总记录的90%左右

源码位置：

https://gitlab.bigdata.leshiren.com/liulian/spark\_mllib\_letv/blob/master/data/handler/RatingIDMaping.scala