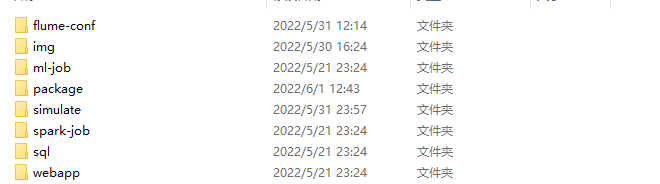
## 系统运行前准备

说明：由于windows下Flume不支持tail -F监控文件以及硬件条件有限，因此

日志采集及备份都在linux下

Spark计算、Redis以及django web服务运行在 windows下



flume-conf flume配置文件

ml-job 机器学习代码

package flume、kafka、zookeeper 安装包

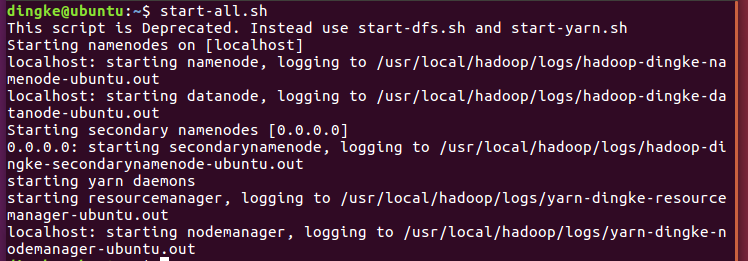
simulate 日志模拟代码

spark-job spark streaming 代码和离线分析spark sql代码

sql mysql表结构及数据

webapp django web 服务

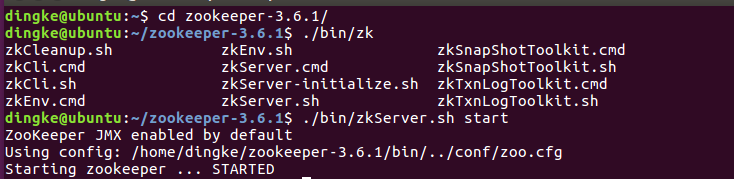
1. 启动hadoop start-all.sh



1. 启动zookeeper

进入zk安装目录下

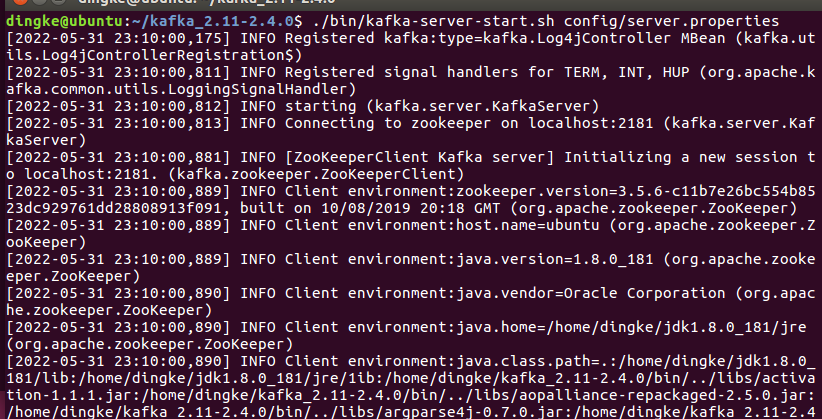
./bin/zkServer.sh start



1. 启动Kafka

进入Kafka安装目录下

./bin/kafka-server-start.sh config/server.properties



创建Kafka主题，进入bin目录下

kafka-topics.sh --create --zookeeper localhost:2181 --replication-factor 1 --partitions 1 --topic weblog-sink-topic

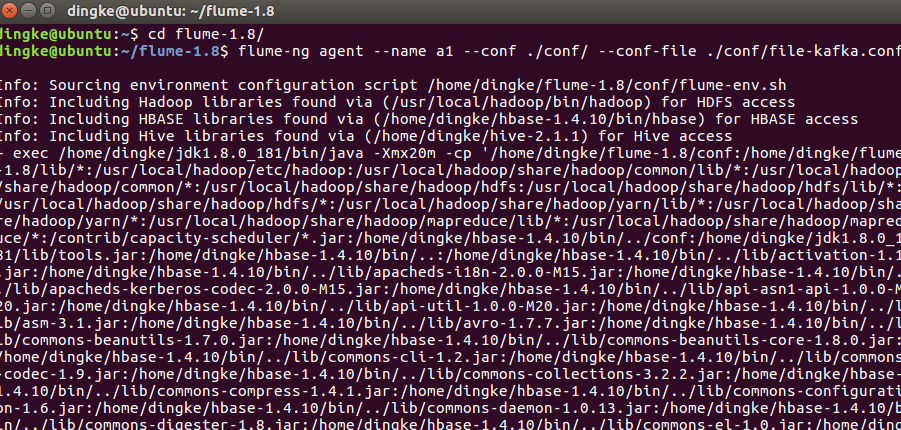
kafka-topics.sh --create --zookeeper localhost:2181 --replication-factor 1 --partitions 1 --topic weblog-spark-topic

1. 启动Flume a1和a3

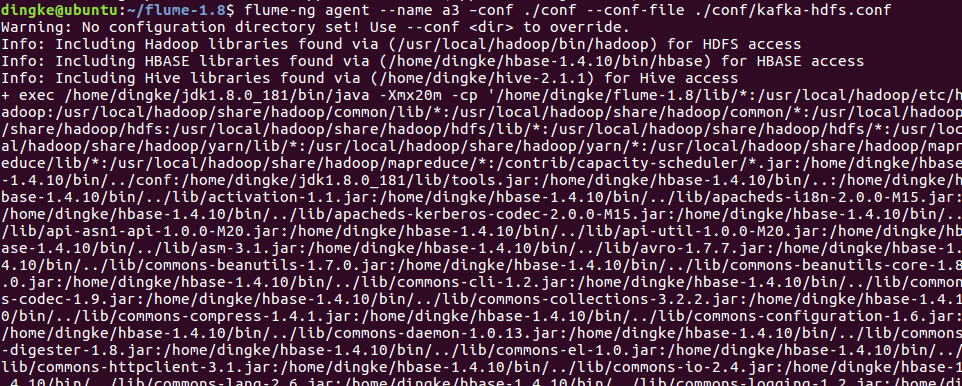
必须要将Flume配置文件file-kafka.conf 、kafka-hdfs.conf 复制到Flume的conf目录下，然后再进行启动

进入flume安装目录下

flume-ng agent --name a1 –conf ./conf --conf-file ./conf/file-kafka.conf



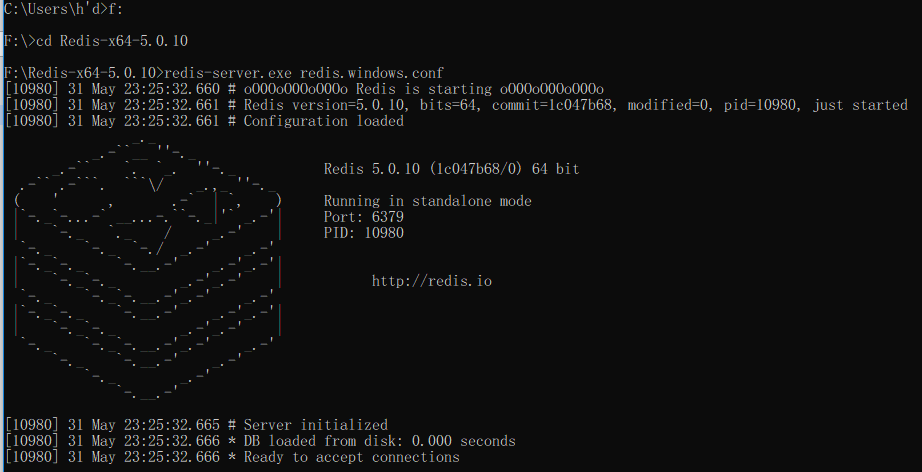
flume-ng agent --name a1 –conf ./conf --conf-file ./conf/kafka-hdfs.conf



1. windows下启动Redis

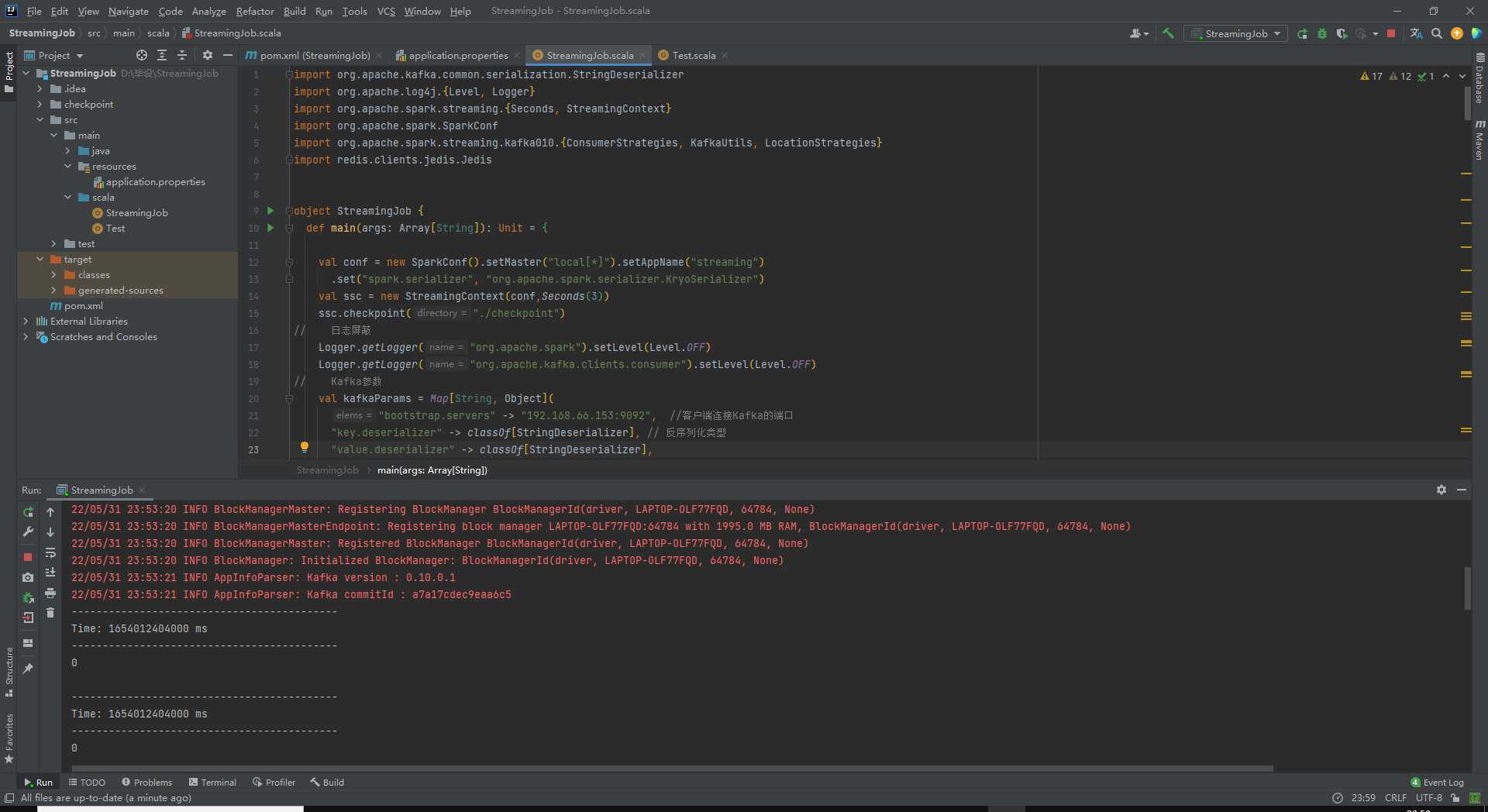
进入Redis安装目录下

redis-server.exe redis.windows.conf



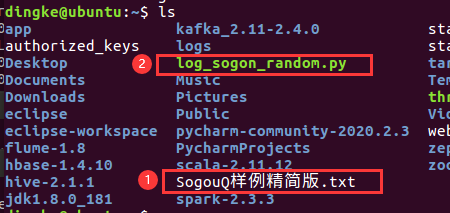
1. windows下启动Spark Streaming

修改bootstrap server端口地址为Linux地址



1. 日志模拟（在实时可视化的时候启动）

将python代码和样例数据同时放在同一文件夹下



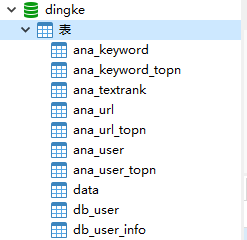
使用python log\_sogon\_random.py –n 模拟数量 进行实时写入



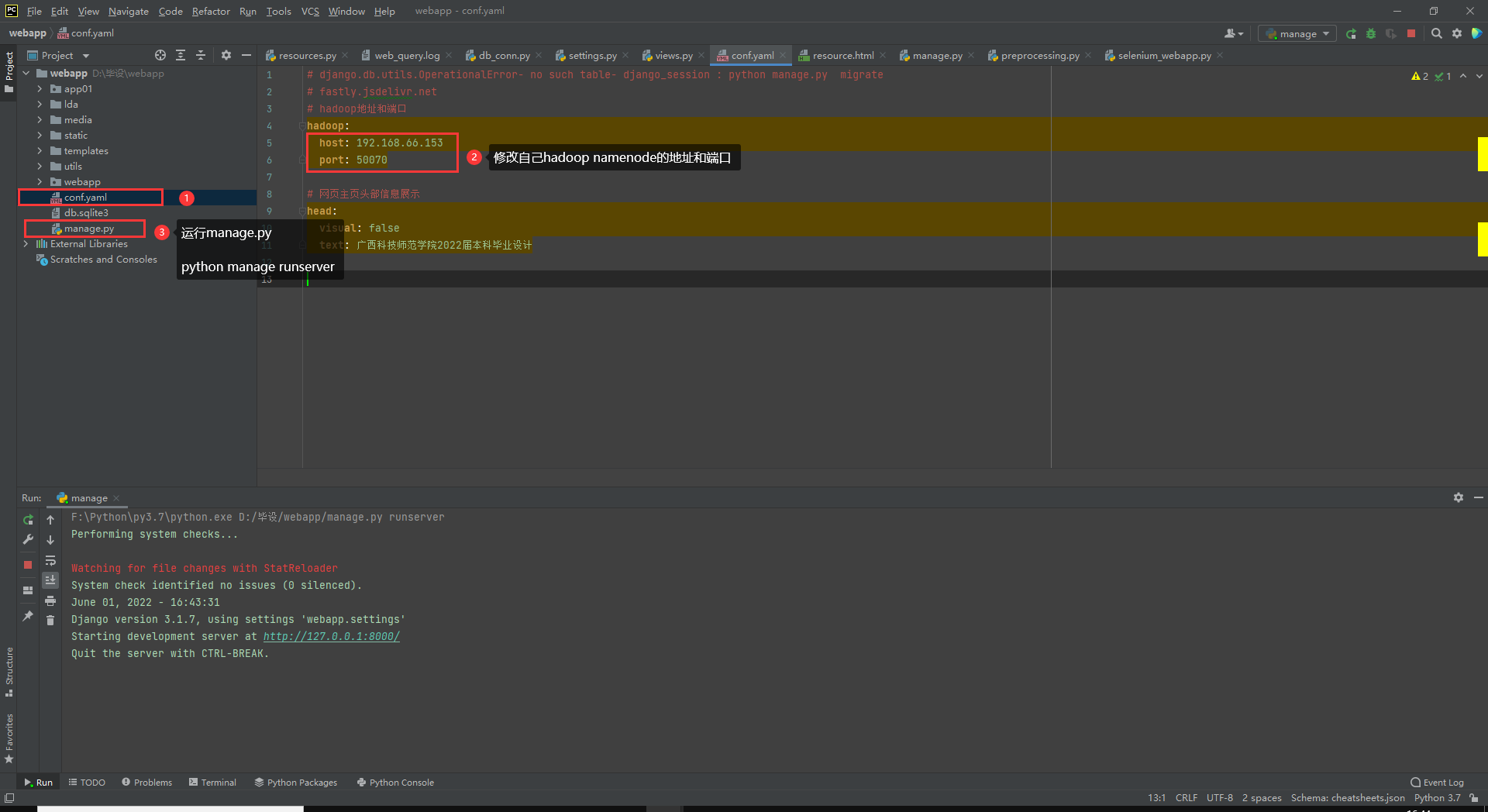
1. MySQL数据库导入

新建一个数据库名为 dingke

导入sql文件，应该有以下表



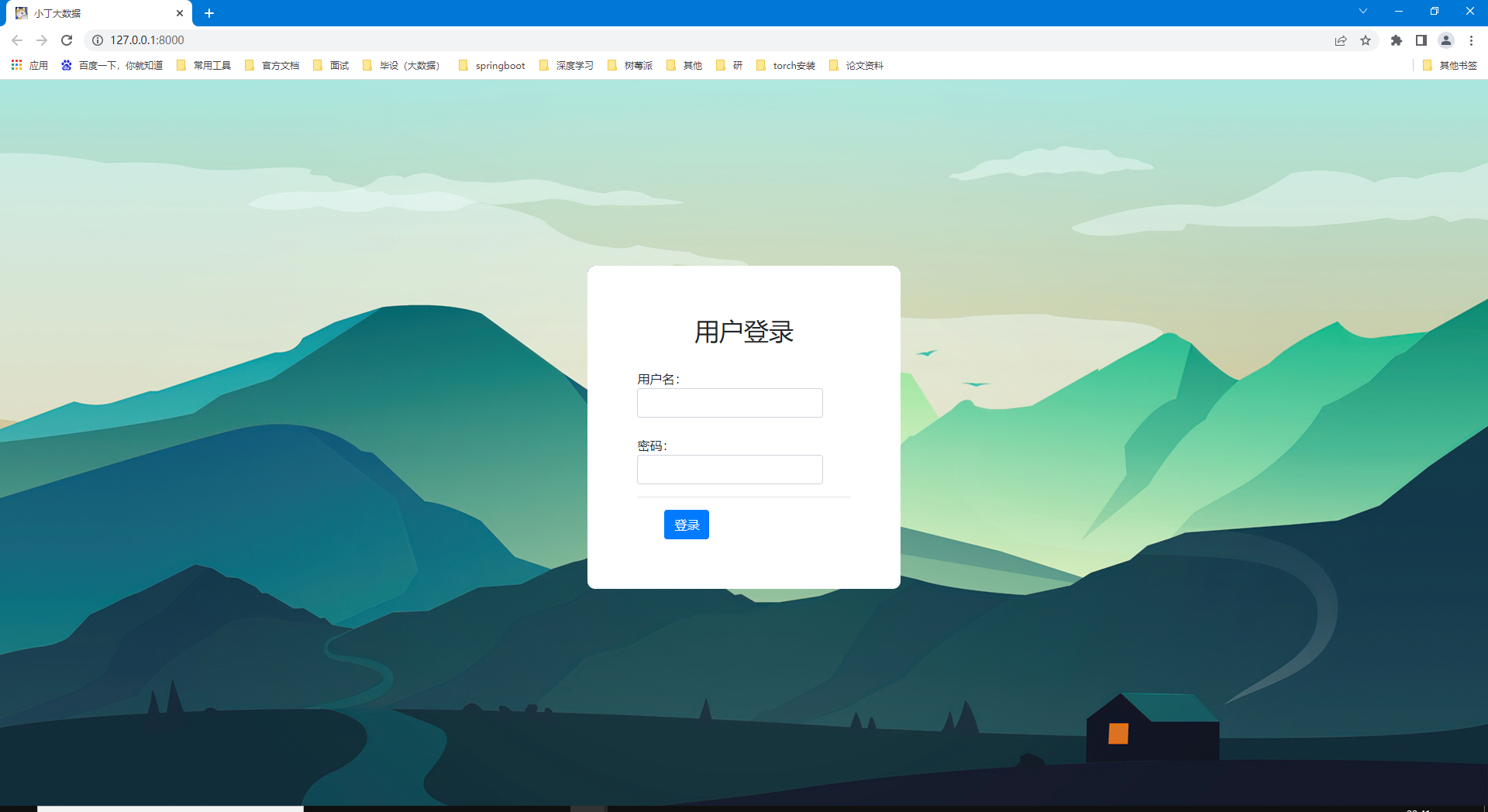
1. 启动django



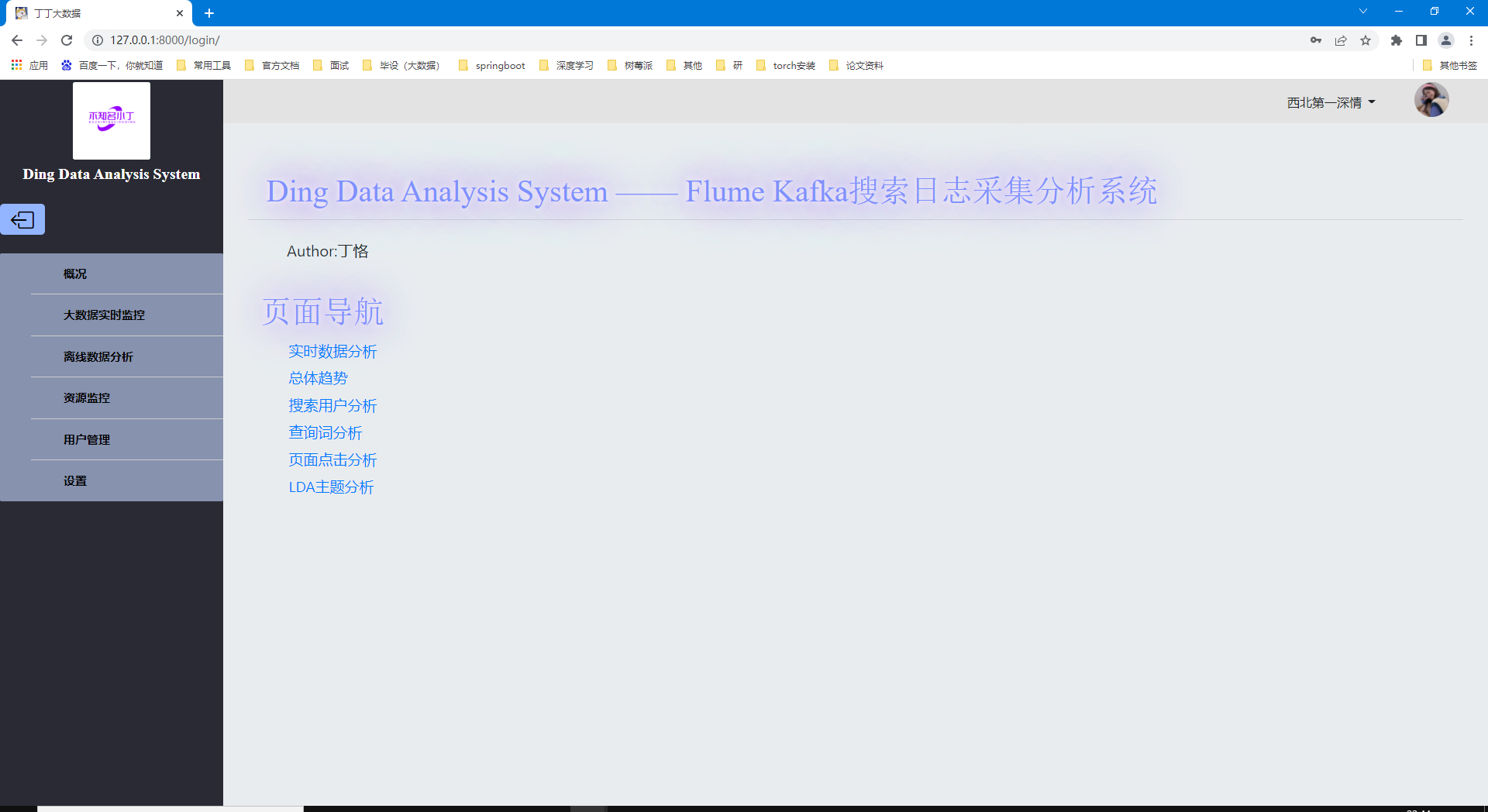
注：演示视频有完整启动过程可以参考

## 系统操作步骤详情

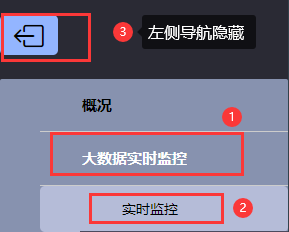
1、运行django项目，打开浏览器，输入127.0.0.1:8000，进入系统登陆界面

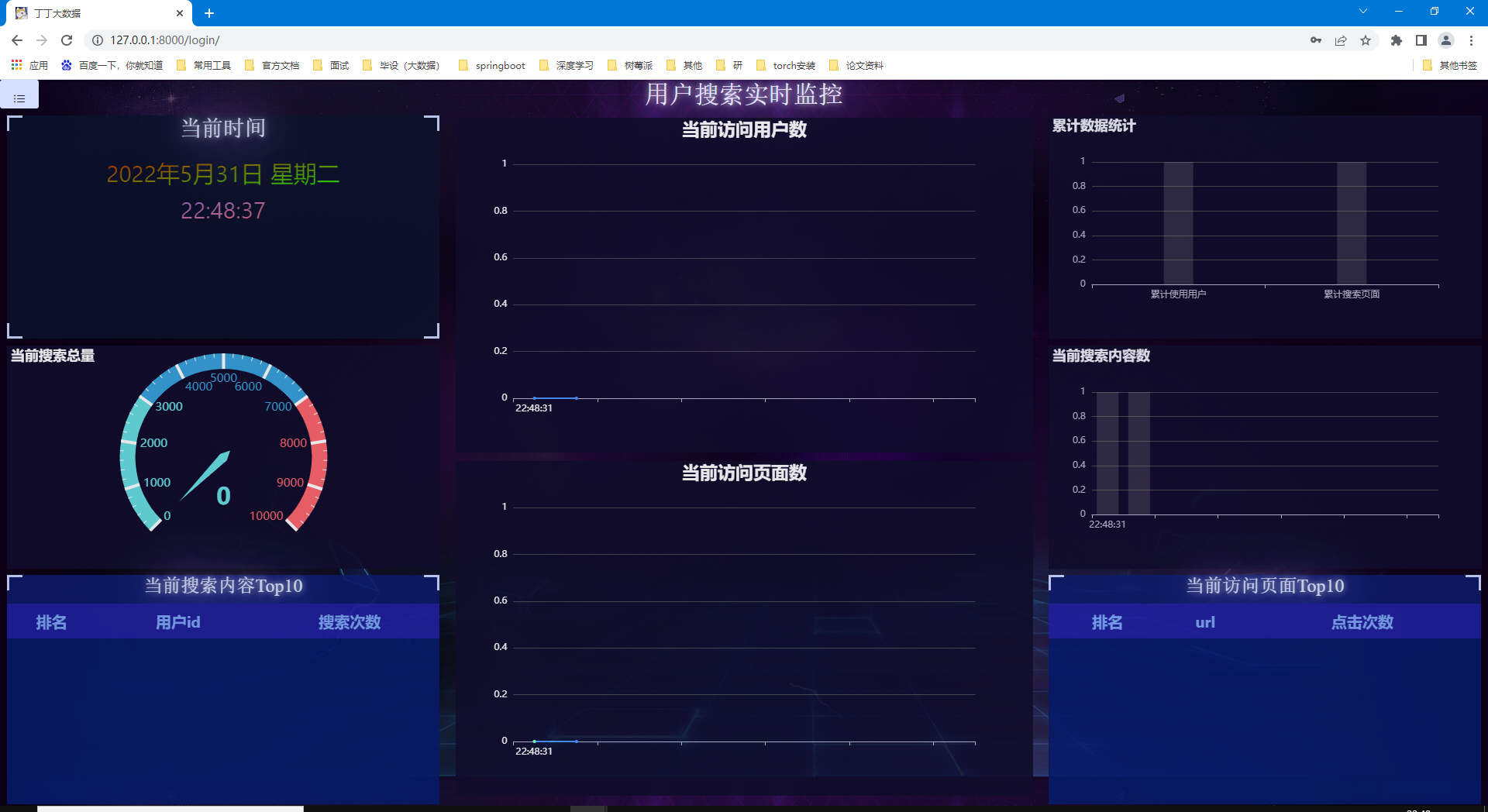


2、输入管理员账号，用户名1646616219006和密码123456点击“登录”进入系统主页



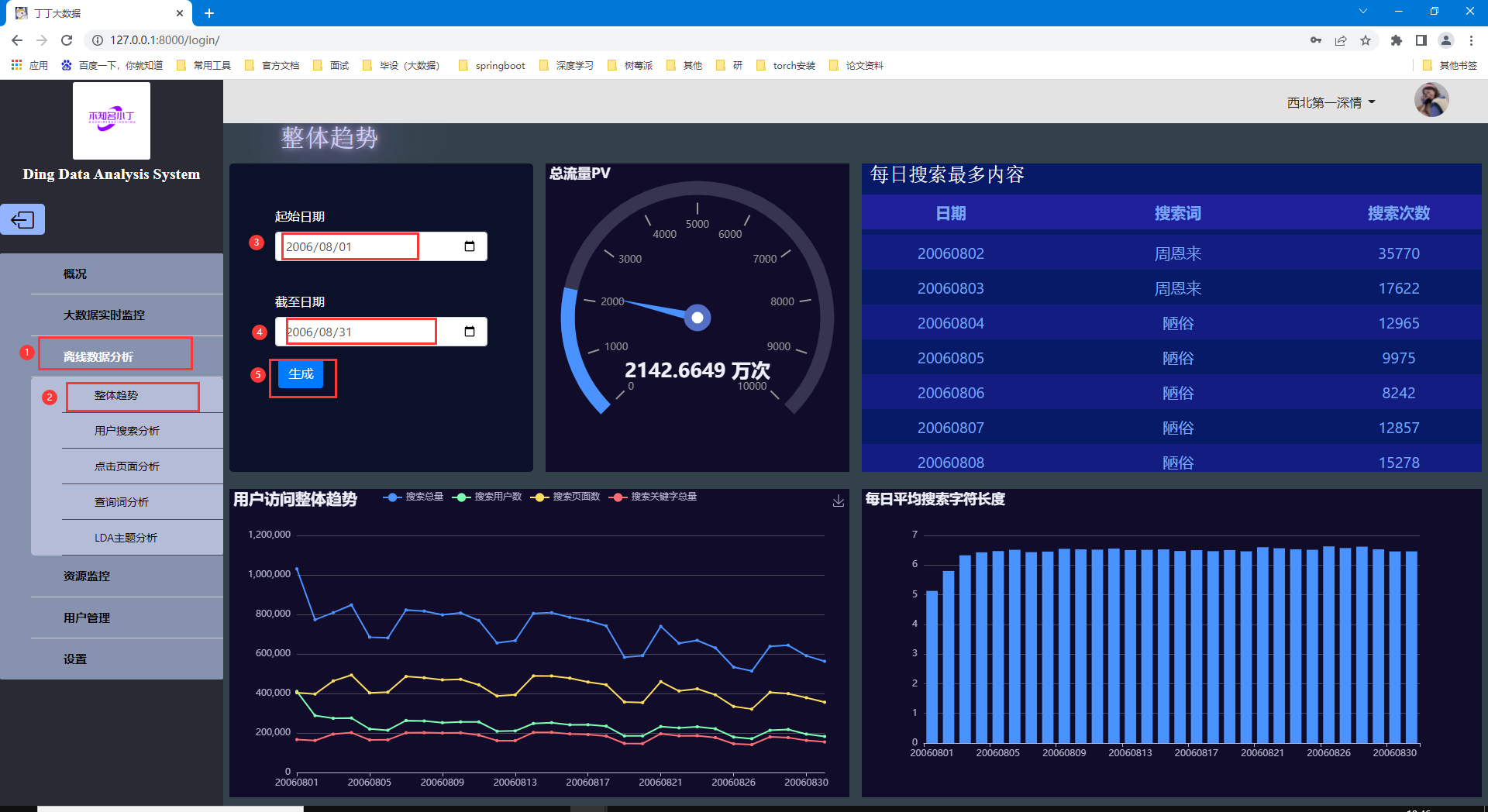
3、按顺序依次点击“大数据实时监控”，“实时监控”，“左侧缩进”即可打开实时可视化大屏



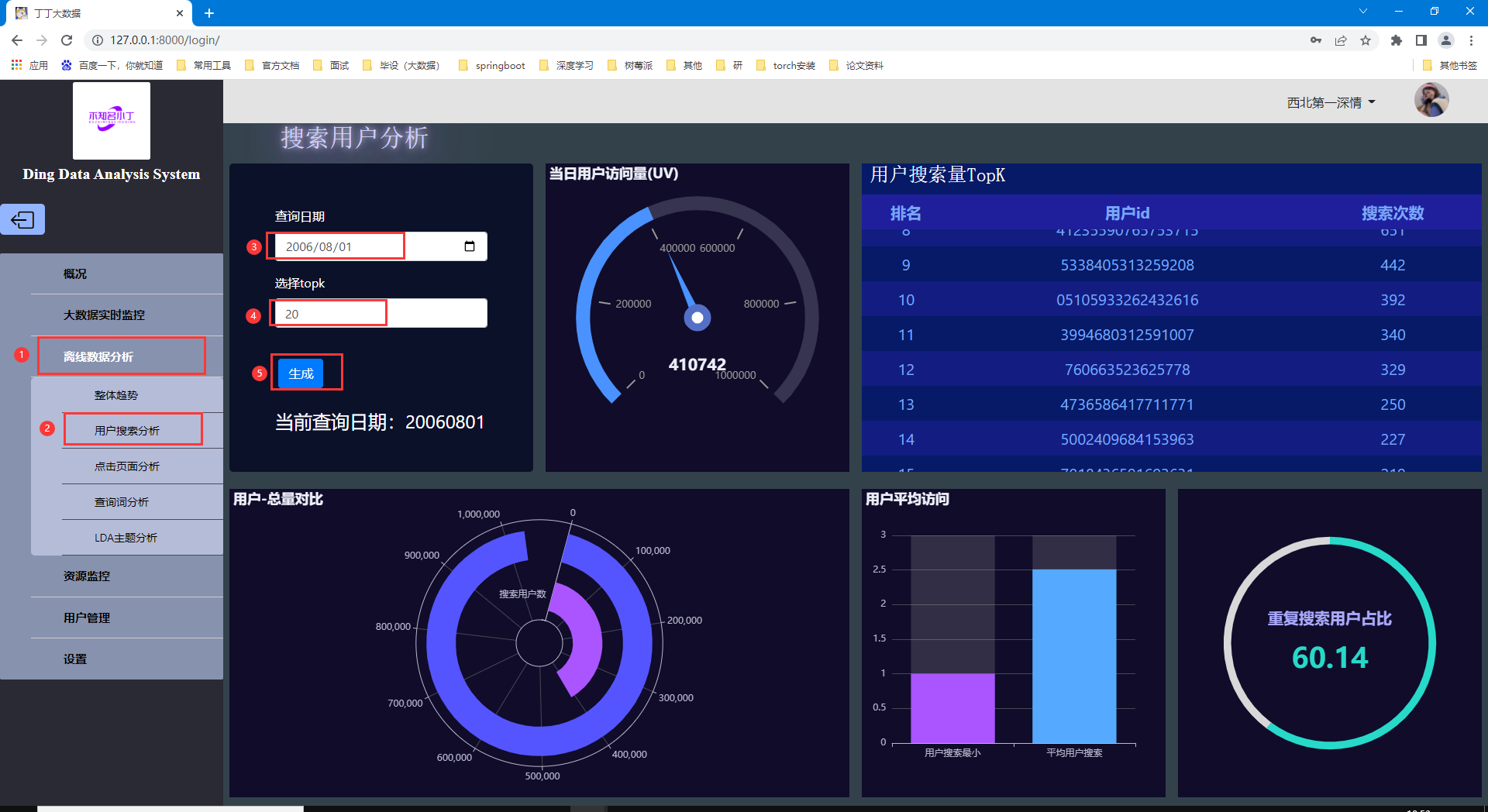


此时页面没有数据，如需实时动态展示，需要模拟日志写入 ，参考如上第7点

4、按顺序依次点击“离线数据分析”，“整体趋势”，选择日志范围，点击“生成”，即可查看指定日期的数据可视化



5、按顺序依次点击“离线数据分析”，“用户搜索分析”，选择指定日期和TopK，点击“生成”，即可查看指定日期的数据可视化



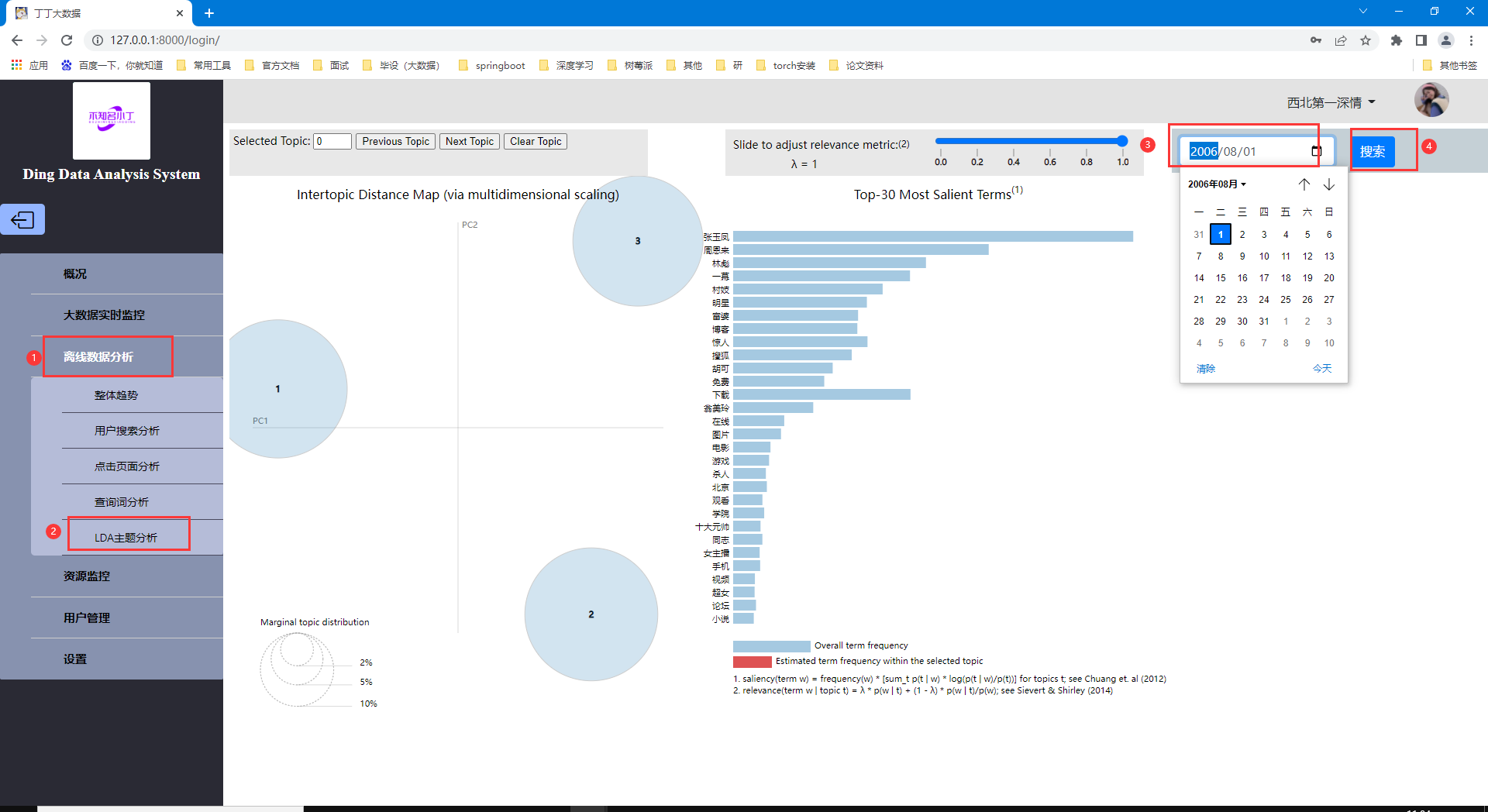
6、按顺序依次点击“离线数据分析”，“点击页面分析”，选择指定日期和TopK，点击“生成”，即可查看指定日期的数据可视化



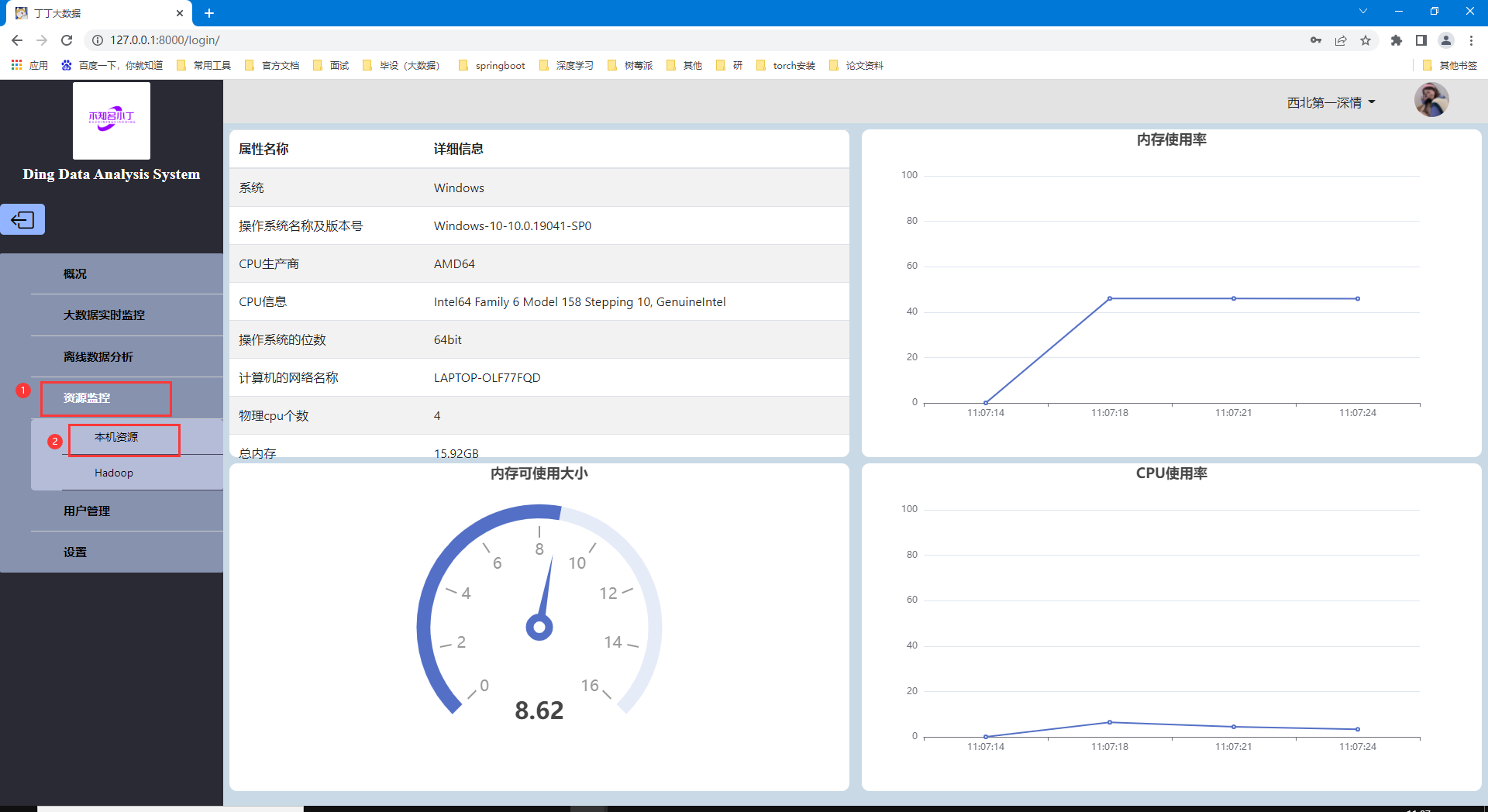
7、按顺序依次点击“离线数据分析”，“查询词分析”，选择指定日期和TopK，点击“生成”，即可查看指定日期的数据可视化



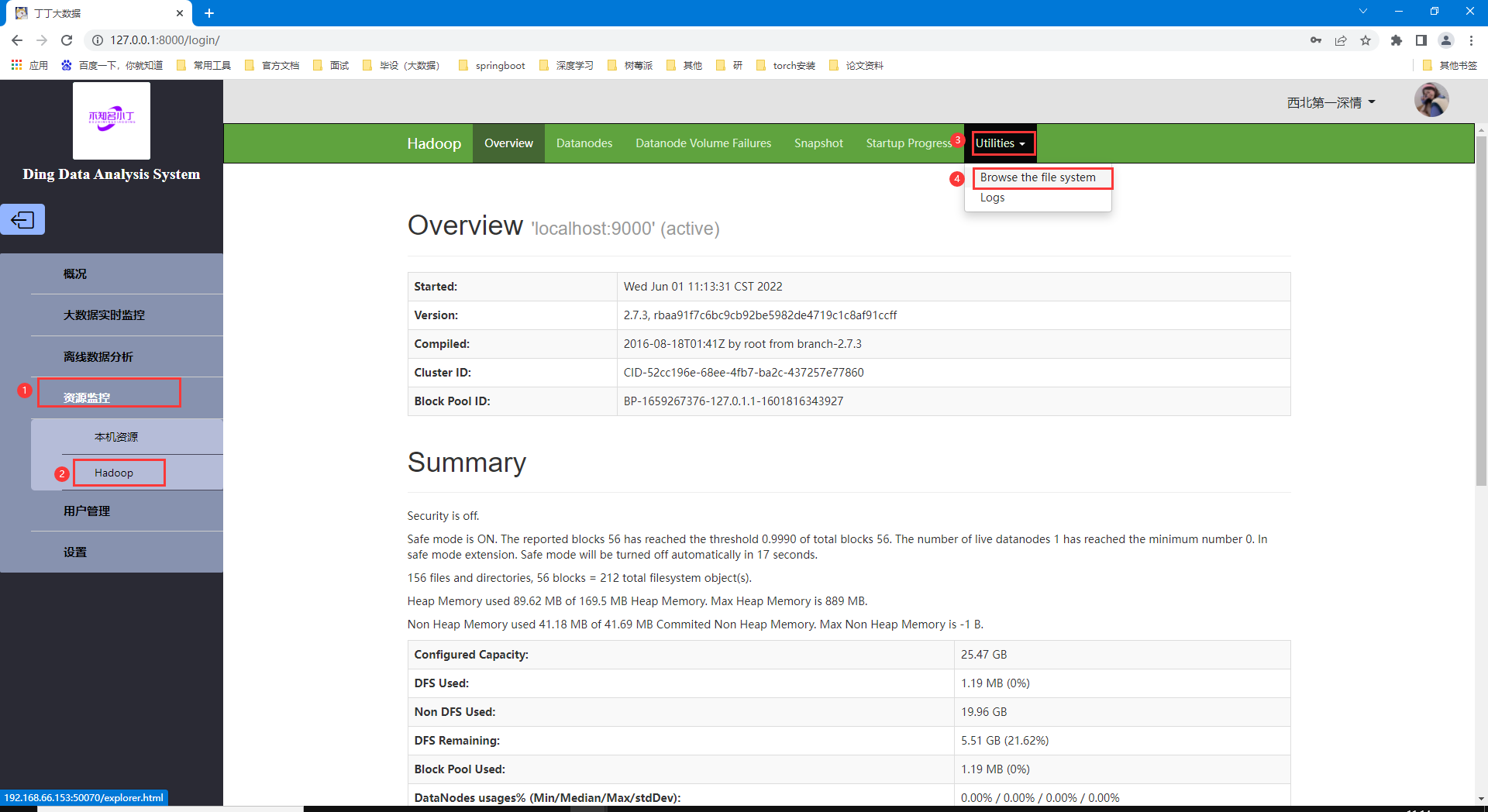
8、按顺序依次点击“离线数据分析”，“LDA主题分析”，选择指定日期，点击“生成”，即可查看指定日期的聚类以及每个类对应的前三十个词

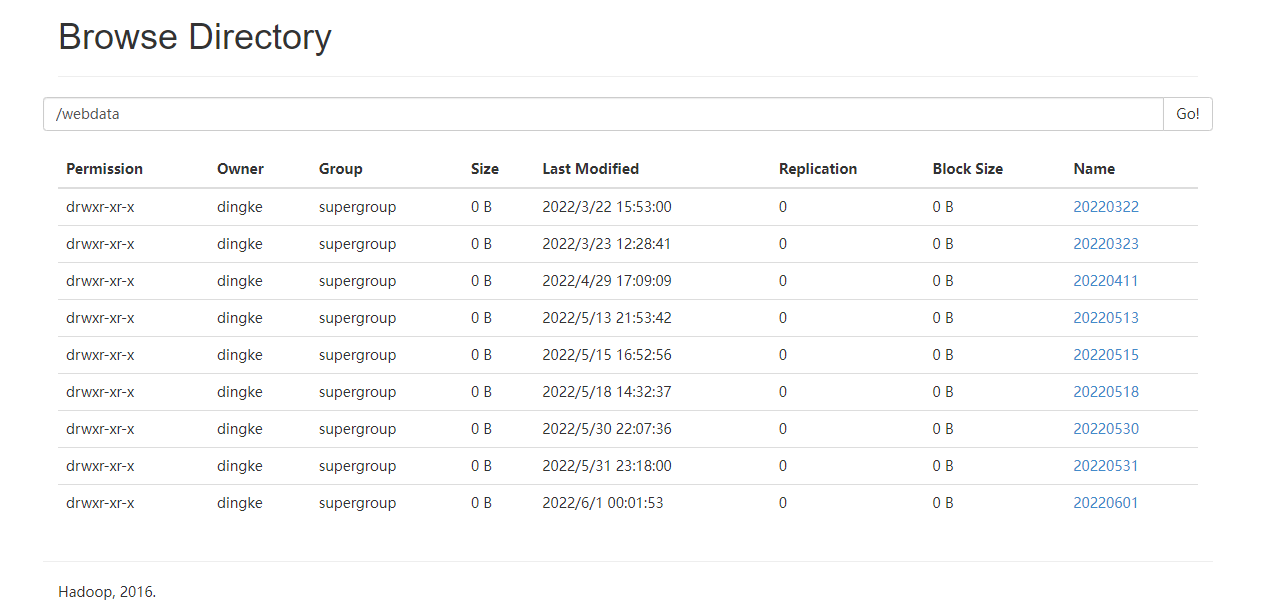


9、按顺序依次点击“资源监控”，“本机资源监控”，即可查看当前本机系统资源消耗情况

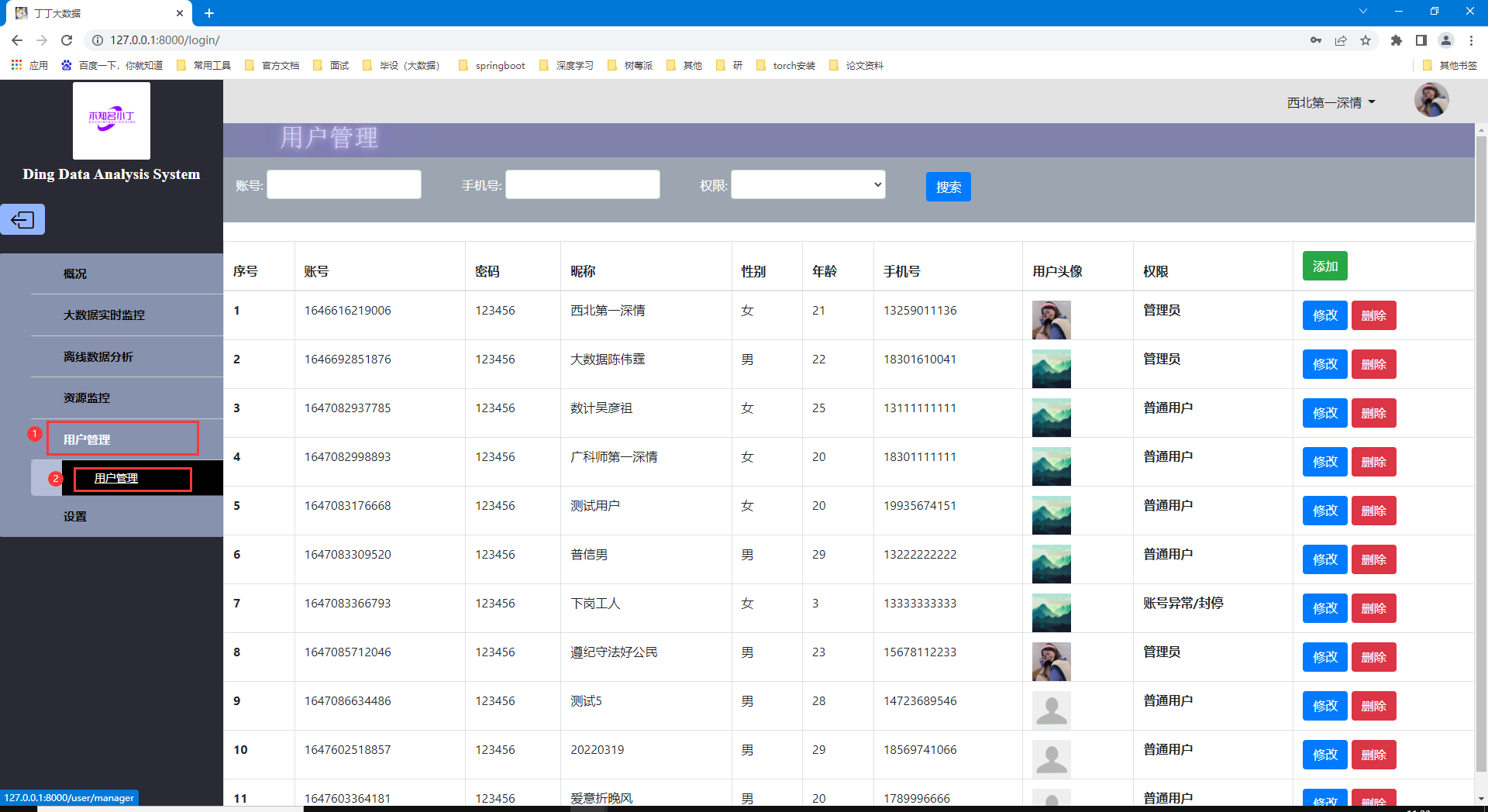


10、按顺序依次点击“资源监控”，“Hadoop”，“utilities”,“Browse the file system”即可查看当采集到HDFS的日志

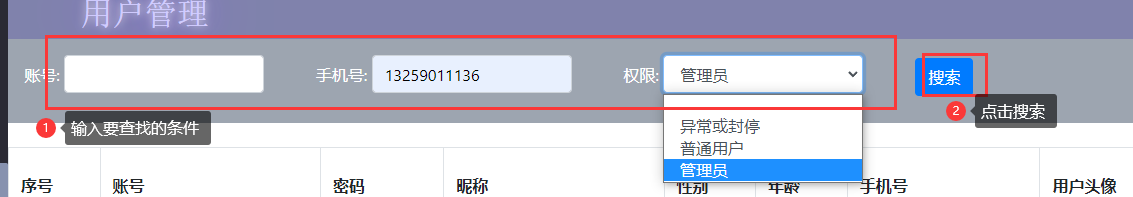




11、按顺序依次点击“用户管理”，“用户管理”即可显示系统用户信息，只有管理员才可以查看，普通用户无法访问

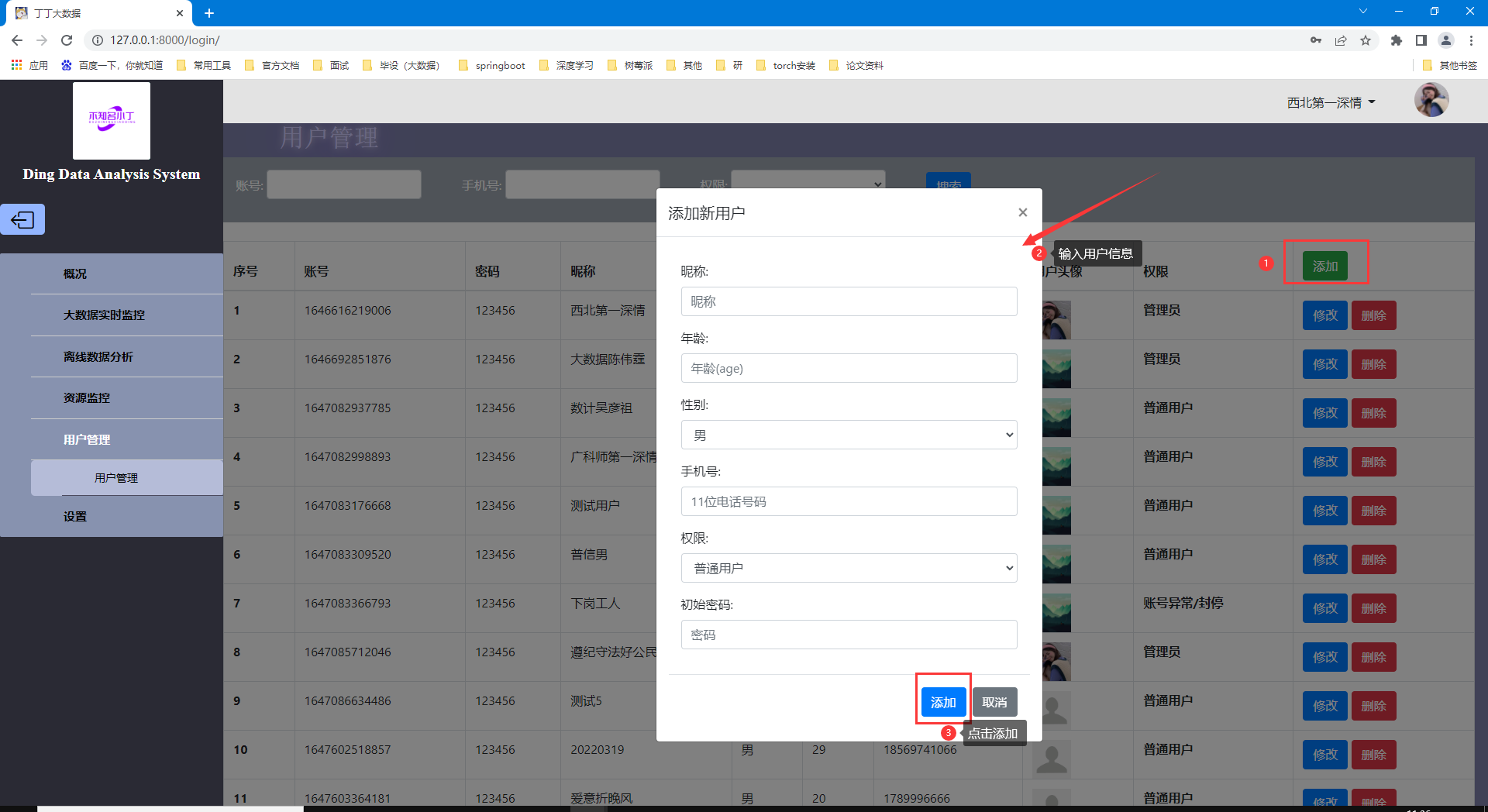


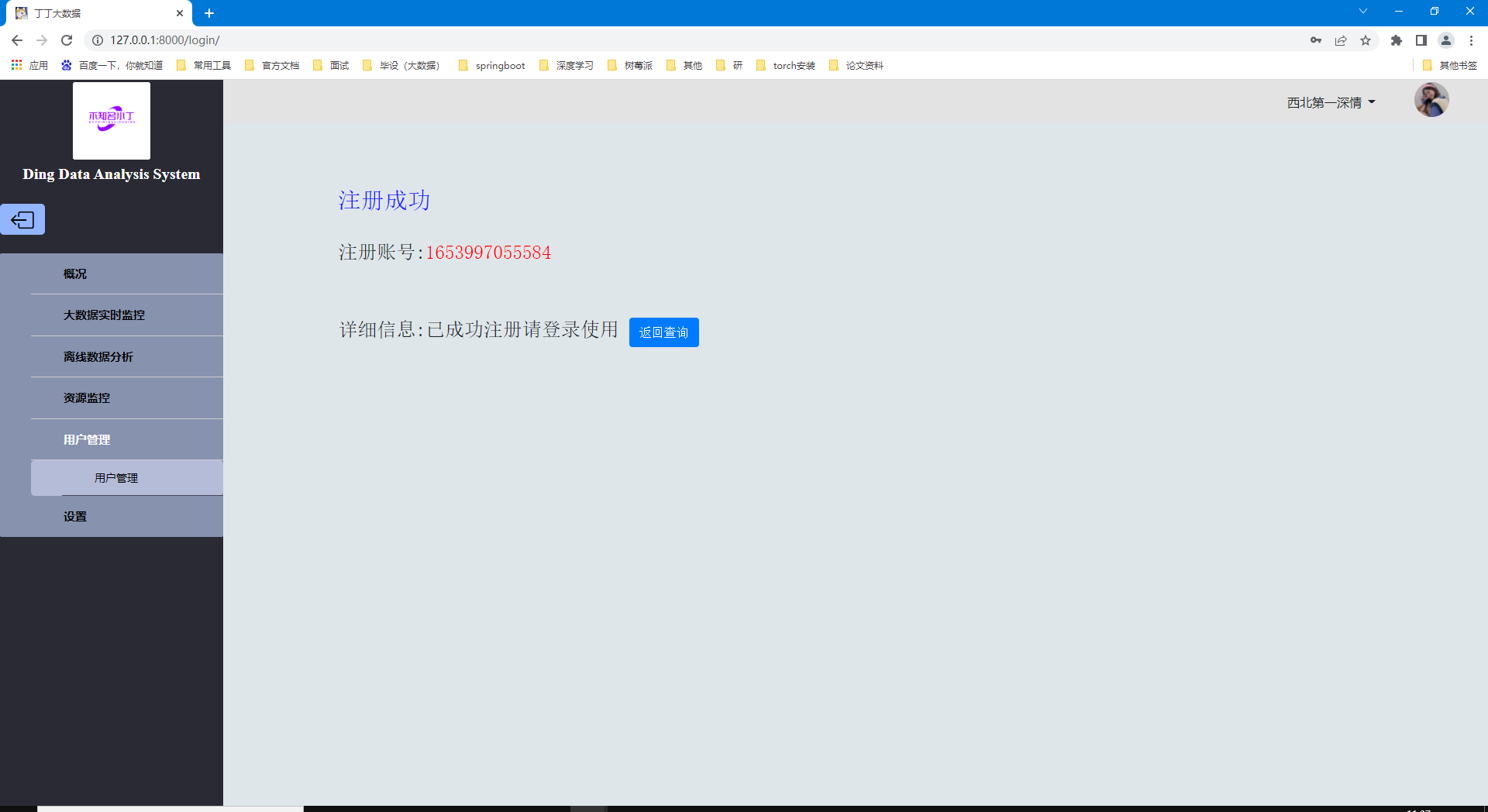
12、在用户管理界面，在条件栏选择要查找的条件，分别是“账号”，“手机号”，“权限”三个条件，可以多条件查询，点击搜索



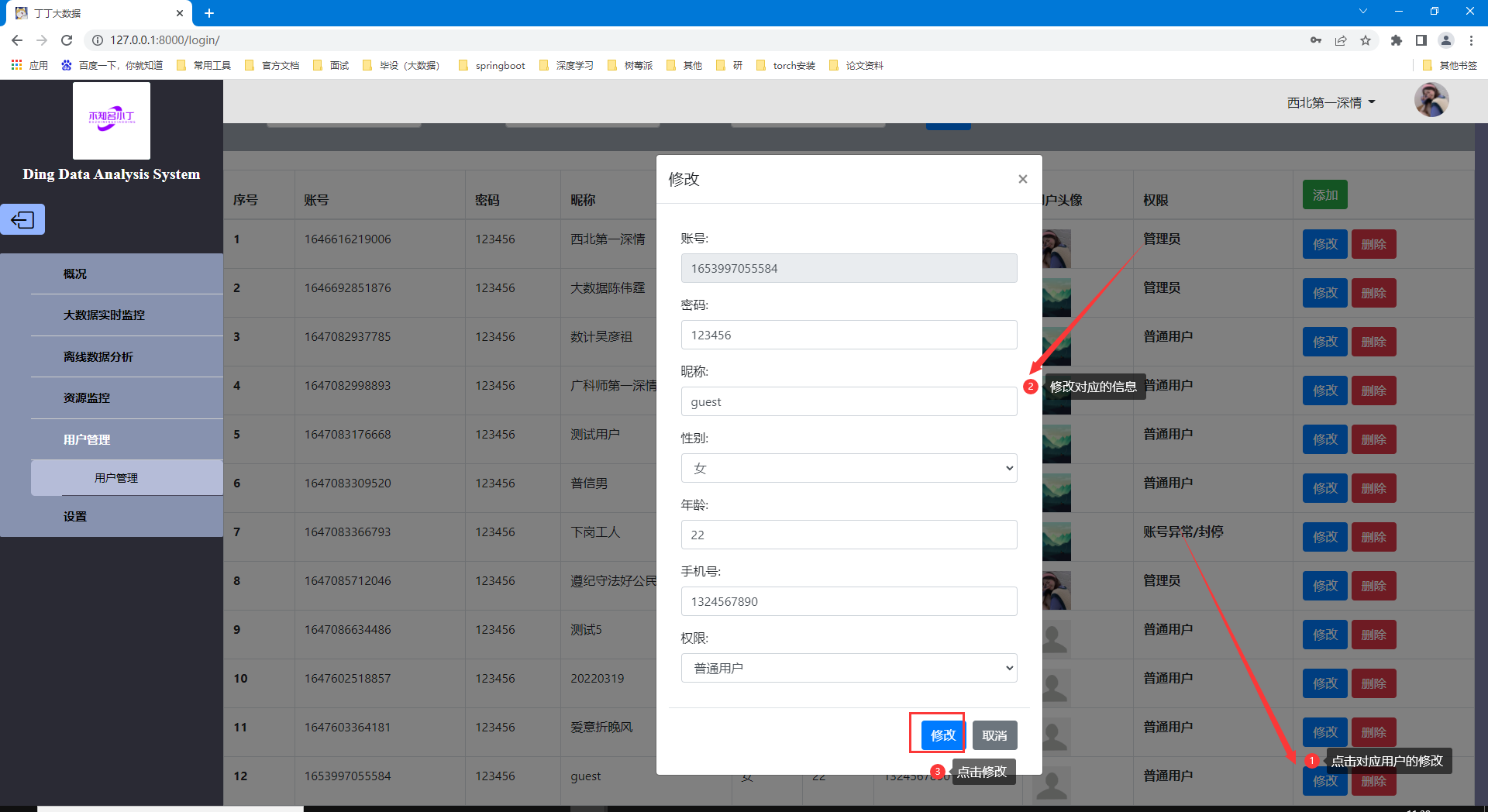


13、在用户管理界面，点击“添加”按钮，输入用户信息，点击“添加”，添加新的用户

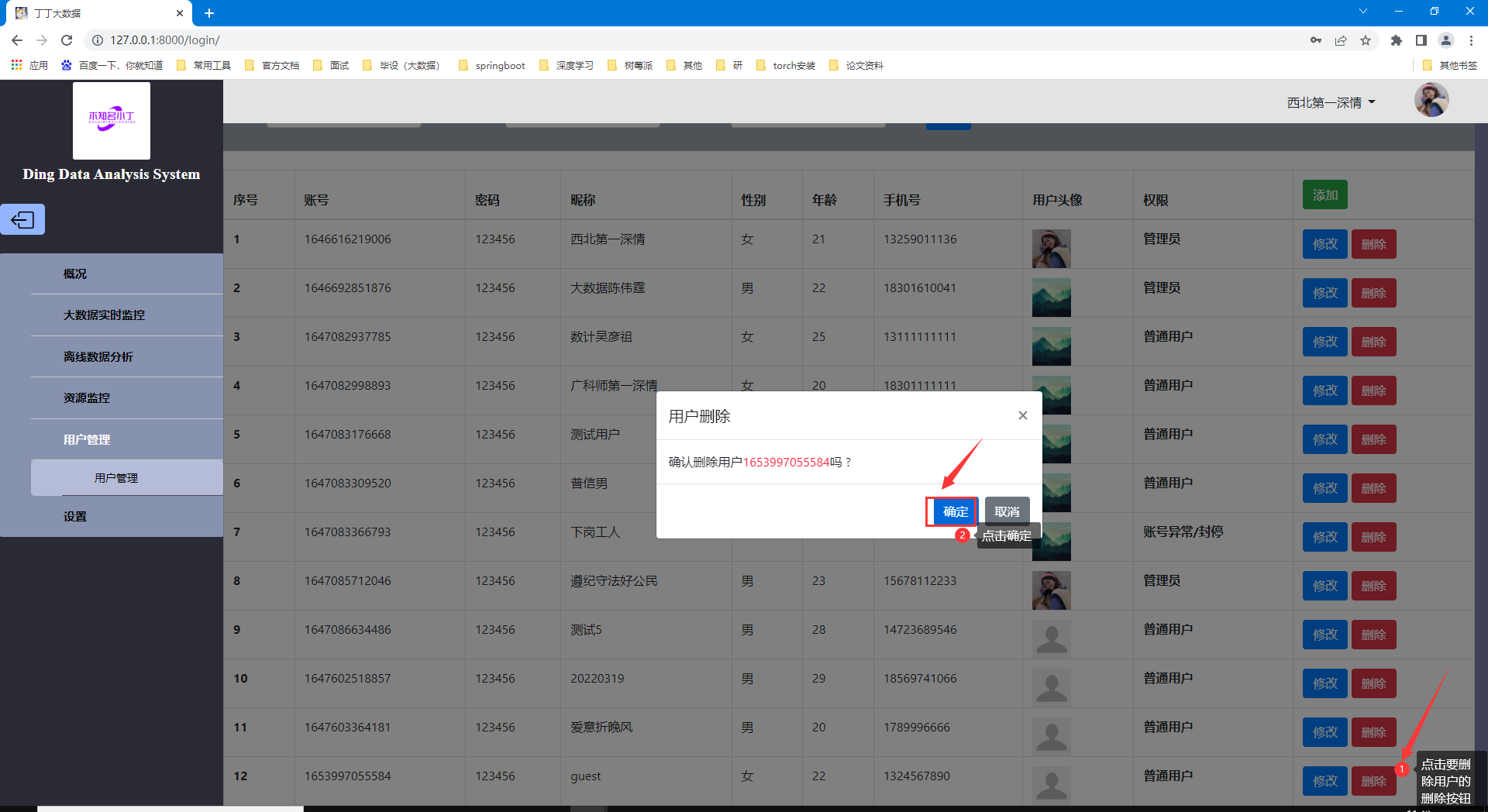




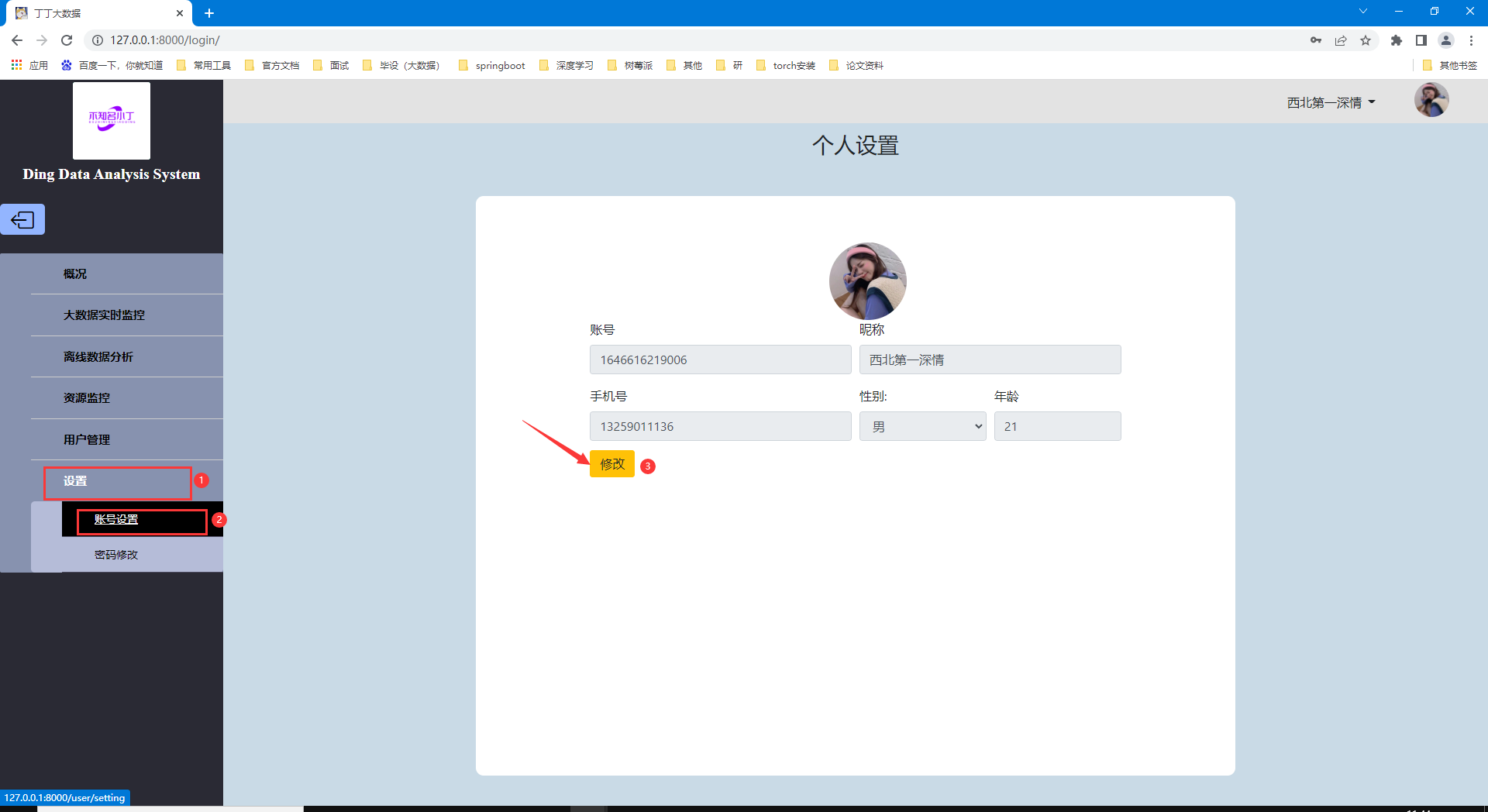
14、在用户管理界面，点击需要修改用户的对应“修改”按钮，修改对应的信息，点击“添加”，完成修改



15、在用户管理界面，点击需要删除用户的对应“删除”按钮，点击“确定”，完成删除



16、按顺序依次点击“设置”，“基本设置”，“修改”用户可以修改个人的基本信息，修改完成点击“提交”后刷新即可





17、按顺序依次点击“设置”，“密码修改”，修改完成点击“确定”后刷新即可

