作品名称：基于日志数据的用户流量预测

1. 作品概述

应用深度学习LSTM等模型，以及结合传统机器学习序列预测用户流量（如每天的访问用户数量）、访问数据类型，从而有效处理网络堵塞、高并发等问题。

另一方面，我们对用户日志进行数据分析，以帮助更好的发现关联特征，统计出经常使用的入口页面。管理者进而优化入口页面，从而获取更好的推广和挽留作用。根据用户搜索的关键词，形成标签云，从而提升用户体验和网站逼格。

作品以网站的形式进行展示。

1. 需求分析

考虑到地质资料馆现有功能服务比如根据经纬度等多条件查询、地图展示等已经相当成熟，但是在用户日志数据分析与处理方面有待提高。我们团队从用户日志数据出发，对用户日志数据进行分析，发掘数据特征并对访问进行预测，从而为管理者和使用者提供更好的服务。

1. 功能设计概述
2. 数据分析

基于用户关键词的分析，基于访问数据类型的分析，基于访问者地理位置的分析，基于访问入口页面的分析。

1. 预测

单日（周、月）用户访问数量预测。

1. 作品运行环境

Python相关环境。可使用内网穿透技术，通过浏览器访问本机从而进行结果展示。

1. 作品制作周期

2018年6月10日-7月1日：收集资料，自学习

2018年7月2日-7月15日：项目编写与测试。

2018年7月15日-7月31日：初稿、修订、项目编码完成。

2018年8月1日-8月15日：整理比赛材料。

1. 参赛宣言

智者创造机会，强者把握机会，弱者等待机会。

1. 团队介绍

队长：李世行，队员：唐壮、高鑫彧、童伟、孙松。

指导老师：陈伟涛、阎继宁。

整个团队都来自中国地质大学（武汉）包括两名从事遥感领域多年的优秀教师、三名在读研究生以及两名即将入学的研究生。



唐壮，就读于中国地质大学（武汉），计算机学院计算机科学与技术专业研究生一年级。在本项目中负责用户的关键词分析，以词云形式形象的进行展示，发掘数据统计特征以及内在关联，从而为管理者和使用者提供更好的服务。



高鑫彧，就读于中国地质大学（武汉）计算机学院研究生一年级。我在这个项目中做的是数据分析，基于访问数据类型的分析，基于访问者地理位置的分析，基于访问入口页面的分析。基于Python语言对大赛提供的数据做数据分析，目标是找出数据的相关性。



童伟，武汉理工大学电子信息工程专业本科生，中国地质大学（武汉）计算机技术专业准研究生。负责部分是应用深度学习LSTM等模型，以及结合传统机器学习序列预测用户流量（如每天的访问用户数量）、访问数据类型，从而有效处理网络堵塞、高并发等问题。



孙松，中国地质大学（武汉）计算机技术专业准研究生。主要配合唐壮，高鑫彧完成数据分析模块的工作。