实训第十次小组例会会议纪要

时间：2018.01.05 19：00~21:00

地点：教学楼S220会议室

组名：灿烂的火星

题目：健身会所顾客精准营销分析系统

参会人员：齐帅彬，刘敏慧，方蕾，黄政峰，陈潇阳

会议记录：方蕾

**会议内容：**

1、组内分享交流

1. 齐帅彬：

基于博士论文《基于协同过滤算法的推荐系统综述》为大家介绍了协同过滤算法的基本概念，对比了基于内存、模型的方法以及基于内容的推荐和混合推进模型四种算法。另外还提出一些难点，主要是数据稀疏性导致的推荐不准确。

2） 刘敏慧：

<http://blog.csdn.net/andrewgb/article/details/50760575>

<http://blog.csdn.net/dark_scope/article/details/17228643>

以以上两篇博客作为本次分享交流的主要内容，介绍基于用户的协同过滤。主要介绍同现矩阵\*评分矩阵的预测作用。

3） 方蕾：

基于《dataminingguide》为大家介绍基于用户及基于物品的协同过滤算法。主要介绍找相似用户的几种方法及其适用范围以及找到相似物品的方法。

* 寻找相似用户的主要方法有：曼哈顿距离、欧几里得距离、皮尔逊相关系数、余弦相似度。如果数据存在“分数膨胀”问题就使用皮尔逊相关系数。如果数据比较“密集”，变量之间基本都存在公有值，且这些距离数据是非常重要的，那就使用欧几里得或曼哈顿距离。如果数据是稀疏的，则使用余弦相似度。
* 寻找相似物品的主要方法是修正余弦相似度。

4） 陈潇阳：

简单介绍了神经网络的基本概念。神经网络是对特征向量做分类，输入是数值向量，输出是分类结果。生物神经网络有很多神经元、突触、细胞体、轴突、树突等，神经元之间有输入输出接口（树突和轴突），通过电信号传递信息。人工神经网络是模拟生物神经网络的。

另外简单介绍了自己的学习计划及进度。

5）黄政峰：提出卷积神经网络主要用于缩减信息量，跟项目关系不大。卷积神经网络可以不用看。

2、 换岗

方蕾：例会记录员🡪进度规划员

陈潇阳：资料保管员🡪例会记录员

刘敏慧：进度规划员🡪资料保管员

黄政峰：进度规划员🡪汇报展示员

3、内部时间大体规划

下周任务：

黄政峰、刘敏慧、齐帅彬负责数据清洗

方蕾、陈潇阳负责网站部分

2.5之前初步分析结果。

3月15日之前把所有工作完成。与后台的衔接部分完成。