**会议纪要 – 2017.1.16**

**每周工作要求：**

* 每周五下班之前将本周所有工作资料（含：工作报告、实验数据、程序、论文、项目资料等）上传到FTP服务器上做备份，所有人员必须严格执行，以防资料因故遗失；
* 每天离开实验室，关闭电脑(包括显示器电源)，关闭显示器电源，关闭电灯，关闭实验室空调，饮水机电源，以保证实验室用电安全;
* 本学期每周需要提交的会议纪要人员：
  + 研究生：李佳、贺星伟
  + 外校研究生：高成亮，朱龙霞，柴瑜晗
  + 外校本科生：暴宇翔（北理工）、贾霏（北邮-国际学院）

**关于纪要的问题：**

1. 会议纪要是对上周工作的总结，是对下周工作的计划；
2. 目前学生会议纪要存在如下突出问题，需要尽快改进：
   1. 总结部分太简略；
   2. 计划部分过于简单，起不到指导下周工作的作用；
   3. 工作内容规划不合理，工作量偏少；
3. 每周五需对前一阶段所有研究工作做备份，在FTP服务器上完成备份。
4. 人走关机，离开实验室务必关闭桌面电脑，显示器电源，以免火灾。
5. **李佳**

* **上周工作**
  + 论文：阅读主题模型和Online Learning结合的文章，确定下一步实验。
  + 项目：1）去华夏银行讨论模型代码；2）完成研发内容；3）模型使用规范文档。
* **本周工作**
  + 论文：用Java实现“Streaming Gibbs Sampling”算法。目标：春节前争取完成第二项研究工作论文的初稿。
  + 银行项目：
    - 根据银信方面的书面反馈，将研发内容2的数据接口重构，尽量和上一期项目中的交易量预测接口一致，重构完成后，发送给银信团队。
    - 修改预测模型的自动更新算法，模型的规范化使用说明。

1. **贺星伟**

* **上周工作**
  + 修改代码的数据处理模块，以及计算了word embedding；
  + 完成模型的tree模块，binarytree 模块，以及attention计算模块。
* **本周工作**
  + 完成代码的损失函数计算模块，完成training 部分；
  + 完成每个模块的梯度更新部分；
  + 完成代码的评价指标MAP和MRR部分，编写test模块。
  + 目标：春节前争取完成论文的初稿。

1. **周建宇**

* **上周工作**
  + 毕设开题答辩
* **本周工作**
  + 工作1：完成文献翻译（至少5000字）
  + 工作2：尽可能完成综述性文献调研，形成科研笔记。**（综合综述调研和深度综述调研）**

1. **朱龙霞**

* **上周工作**
  + 与李佳沟通LDA和BTM的相关算法基本内容，及实现流程。
  + 实现用LDA算法对英文语料库的处理
  + 学习ICTCLAS的使用及相关信息
* **本周工作**
  + 与李佳沟通BTM：详细了解BTM的实现流程及注意事项
  + 能够用BTM对英文语料库处理：注意对英文语料库处理的步骤
  + 能够用BTM对中文语料库处理：注意在对中文语料库处理与英文语料库处理有什么不同，并提出自己的想法。
  + 在1月22日报告总结会上，总结上周和这周工作。

1. **柴瑜晗**

* **上周工作**
  + 论文：Topical key extraction from folksonomy through graph-based ranking.这篇论文主要做概念抽取相关工作。

第一步是进行主题分布。所用到的工具：利用LDA主题模型进行主题分布。

第二步为标签排序。所用到的工具：PageRank(PR)、Topic-Sensitive PageRank(TSPR)、Topical PageRank(TPR)、context-sensitive topical PageRank(cTPR)、topic-sensitive tag ranking(TSTR)

* + 项目：在python环境下学习LDA工具以及PageRank工具。
* **本周工作**
  + **论文：**看近五年知识图谱相关论文（综述性文献、每一步的具体性文章），进行调研工作，整理思路。按照每一步，以及每一步的代表性方法，并将代表性方法做成表格形式，表格中包括方法名称、方法思路、优点和缺点。
  + **注意：**调研的思路是综合性综述文献调研 和 深度性综述文献调研相结合。

1. **高成亮**

* **本周工作**
  + 熟悉和彻底掌握王九硕的情绪分类算法和代码思想
  + 在本地机器上能够执行算法代码
  + 形成对新的实验思路的尝试