**杨长沛**

论文：基于风险短语自动抽取的上市公司风险识别方法及可视化研究

1.可考虑的点：对于信息产业，公司和国家间，公司与公司之间的比较，公司 自己不同时期的比较

2.如何理解词云图

3.下一个阶段的目标：论文阅读数增加，有自己的框架，讲自己的成果

4.一些key words:主题提取，文本抽取，实体抽取

5.推荐的书：《知识图谱和深度学习》

**谢树松**

论文：Bert

1.bert出现的本质原因是？自bert出现以后，它的出现基本干倒了前面的所有方法，目前很多研究都在围绕bert展开。

2.为什么bert需要设置MLM和NSP这两个任务？

3.与bert模型有关的应用：信息检索，主题抽取

4.聚焦：1）自己构建数据集，人工标注关于稀土方面的专利，应用bert模型， 做应用创新。

2）用公开的数据集，做bert模型方面的创新。

**张文豪**

论文：融合多层注意力的方面级情感分析模型

1.方面信息是指什么？

2.每一层架构的输入输出是指什么？

3.数据集是怎样的？预处理应如何进行？

**姚秀:**

论文：基于专利数据的企业技术创新机会识别方法研究

1.代码、模型、算法和程序需要多花时间

2.很多比赛可以多多参与，比如阿里云 天池大赛

**南康龙盛家具的一个项目**

1. 沙发的工艺流程，采集数据
2. 用智能优化算法对工艺流程做优化

关于看论文的建议

* 这论文为什么这么写？
* 为什么能发表？
* 创新点在哪？
* 我的创新点在哪？
* 尽力看有源码的，方便复现