**张燕琴**

论文：Identifying FinTech Innovations Using BERT

1.现在的分类很多都基于人工标注，可以考虑人工+自动抽取，或者语义分析。

2.专利的价值、创新的质量、提料专利、专利相似度、新颖度，多看论文。

**姜炫至**

论文：词和短语的分布式表示及其组合性  
1.公式推导需体系化，继续深挖这个模型，多看论文。  
2.学习概率统计。

**李奎**

论文：基于深度学习的语言模型研究趋势展望  
1.深度学习结合语言模型有很多方面内容都可以做，需要聚焦，找到方向。  
2.在bert基础上可以增加很多内容（bert+）。  
3.现在有对词、短语的研究，短语包含语义，现在很多也还是在研究词。  
4.很多研究都是基于统计，没有语义,知识图谱是包含语义的。  
  
**夏卫欢**

论文：面向自然语言处理的预训练技术研究综述1**.**bert不能训练到真正的语义信息。

2.找到研究的点，搞明白模型。  
  
**余廷杰**

论文：模板驱动的神经机器翻译  
1.要有程序，尽量可复现论文，并且在基础上进行修改。  
2.可考虑的点：中英文的翻译比较，相似  
  
**邹振远**  
1.继续看代码，公式推导

**其他**

聚焦bert之前的模型代码，进行公式推导，分工理顺框架，使大家有共同的基础，找出点，方便讨论和提问。  
  
**要求**  
现在每个人都要开始准备研究生创新基金项目，便于开题报告。

大方向有个分工，比如数学能力好一点的包括概率统计等、代码、idea多的，例如模型组合等。