信息科学与工程学院毕业设计(论文)

中期检查报告

**题 目**

**学 生**

**指导教师**

**专 业**

**班 级**

**日 期**

**信息科学与工程学院**

|  |  |  |
| --- | --- | --- |
| 1. 工作的主要进展和结果（附相关材料） 2. 了解了CNN/LSTM的主要工作机制，并且对现有的比较有效果的CNN模型框架进行了筛选，考虑到训练的效率以及训练机器的性能，考虑采用VGG-16模型进行CNN层的训练；为了能够发掘局部数据与整体数据的关系，拟采用时间窗的方法对数据进行切分，考虑到时间窗包括前后的时间切片，为了发掘前后时间切片的数据对于整个推荐效果的影响，拟采用LSTM方法对CNN整合后的特征图进行处理； 3. 考虑到不同时间窗口下数据稀疏程度的差异，我们将整个数据矩阵利用embedding方法翻译成两个矩阵（Positive和Negetive），其中Positive矩阵在concat空白处补一个近似与0的数（在具体实现中补0.5），negetive在空白处补0，这样可以在训练过程中模拟出数据在稀疏程度上的差异性变化，从而使得模型更具有鲁棒性； 4. 在对embedding矩阵进行处理时，我们采用re-embedding的方法将Positive/Negetive矩阵进行再次翻译，得到的Q矩阵（对应User维度）,K矩阵（对应Item维度），V矩阵（对应Time维度）进行CNN训练，目的在于发现Q,K,V三个矩阵当中的局部特征信息（因为数据集是排好序的）。 5. 对任务书的要求和计划作了哪些必要的调整和变动   在设计attention层时，我们采用CNN+LSTM的方法来拟合任务书中的Activation Unit，并且在数据维度上，我们选用了UIT三个维度来取代任务书中的UI两个维度 | | |
| 1. 哪些内容未按计划进行，原因何在 2. 训练效果不明显；由于单个VGG模型有1亿3千万个参数，因此在训练过程中产生的梯度消失的问题非常严重，虽然我们在代码中加入了dropout降低运算复杂度，但是梯度通常在Conv3层消失；   2、数据量不足；由于我们的模型训练过程是一个包含了三个VGG和两个LSTM的训练过程，因此单个MovieLen不能满足我们的训练需求，必须找到更大的数据集。   1. 下一步工作预测及存在问题 2. 为了解决由于神经网络的庞大体量造成的梯度消失而无法训练的问题，我们将在之后的研究过程中引入迁移学习的思想，即冻结VGG除Flatten层外的所有层，利用当前的数据对Flatten层进行重新训练，保证我们现在有的数据能够准确适配到模型当中； 3. 目的参数的调节工作，如Positive矩阵中的补全值0.5，可以在之后的工作当中加入cross-validation的方法，对补全值进行暴力搜索，以达到一个比较好的补全效果。 | | |
| 检查结论 | 指导教师意见 |  |
| 检查小组意见 |  |