

The Course Report of Artificial Intelligence

**SCHOOL:** SCHOOL OF SOFTWARE ENGINEERING

**SUBJECT:**SOFTWARE ENGINEERING

指导老师： 宋恒杰

学号： 201530611999

姓名： 李可欣

班级： 15级3班

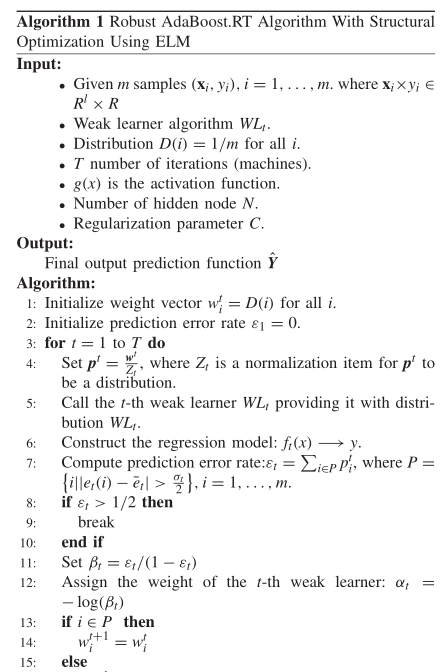
日期： 2018/1/18

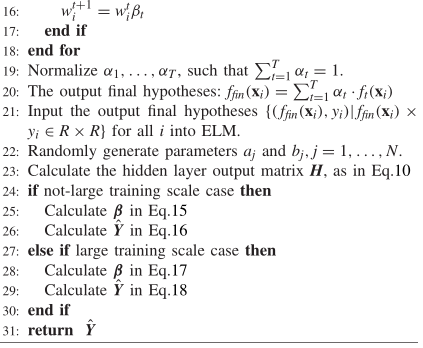
1. **论文原文**

Peng-Bo Zhang and Zhi-Xin Yang,“A Novel AdaBoost Framework With Robust Threshold and Structural Optimization，”in *IEEE Trans. Cybernetics*, vol. 48, no. 1, Jan. 2018. (原文见AdaBoost.pdf)

1. **翻译**

论文翻译见，翻译.pdf。

1. **算法设计及实现**
2. **算法设计**

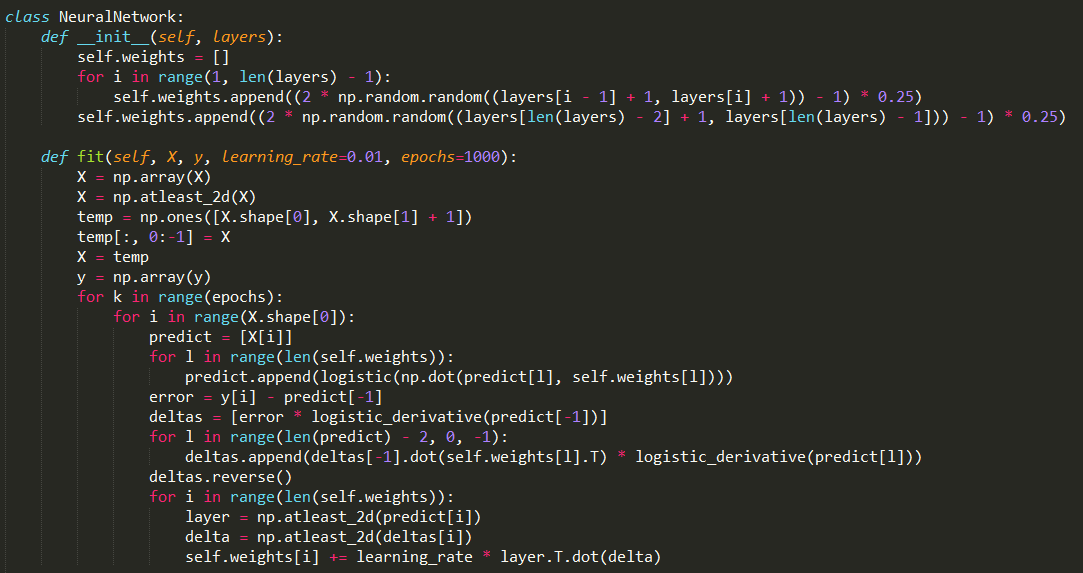


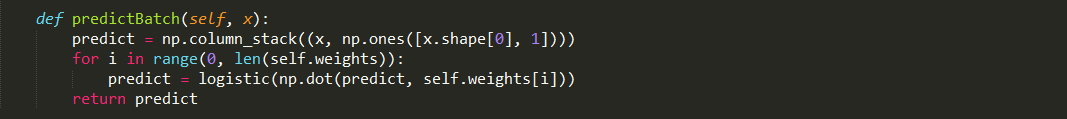
1. **实现**

此次课程任务中，一共实现了包括Back Propagation，AdaBoost，Extreme Learning Machine三个核心算法。其中BP算法作为AdaBoost的弱学习器；ELM作为进一步优化AdaBoost结构的优化方法。实现语言为Python。

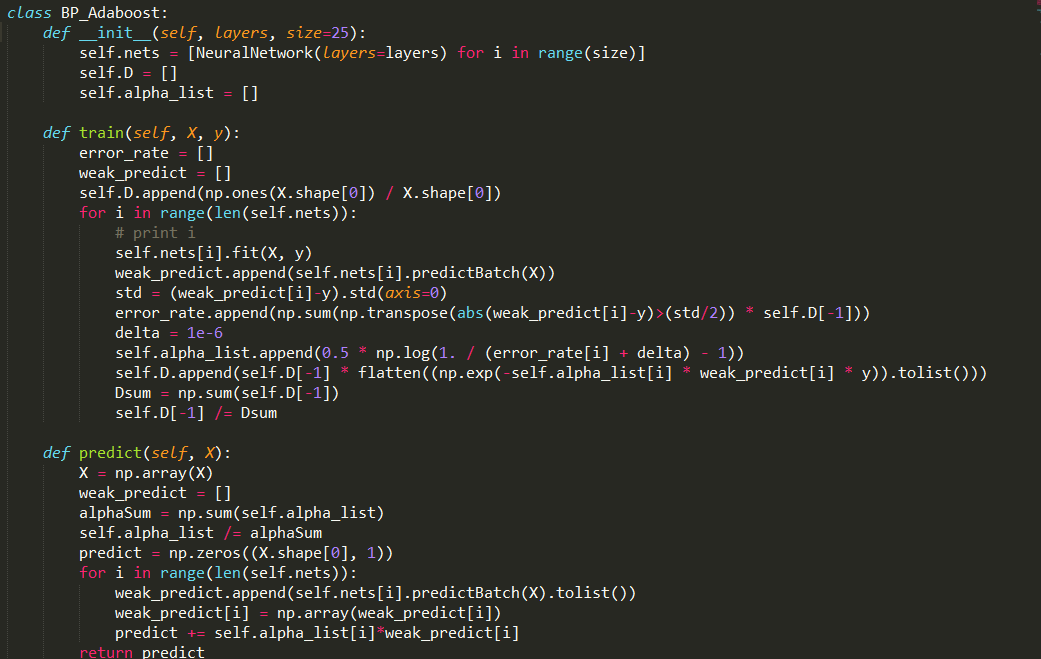
算法实现主要代码截图：

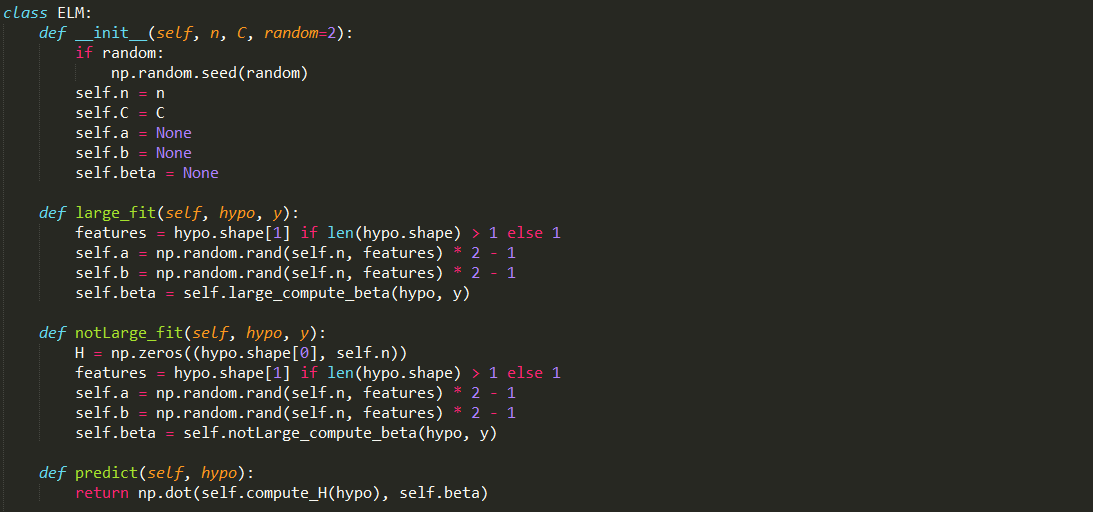
（详见BP.py，AdaBoost.py，ELM.py）

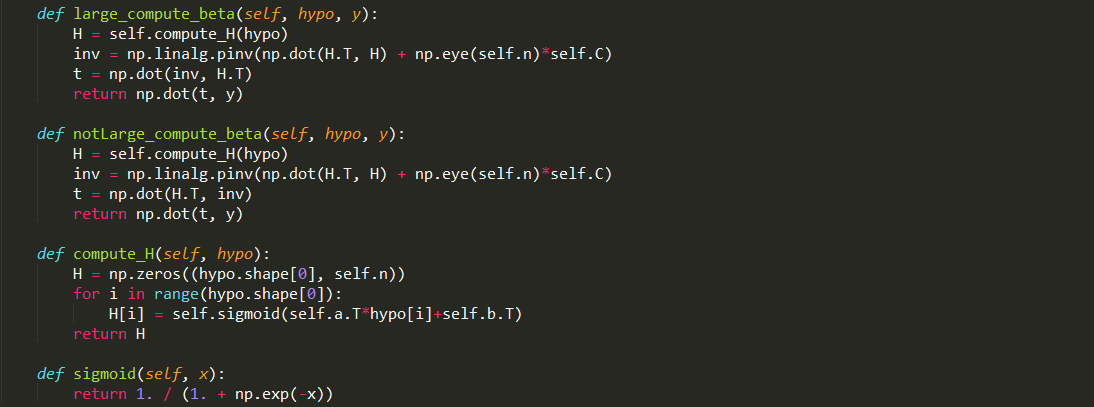
* **Back Propagation**



* **AdaBoost**



* **Extreme Learning Machine**



**四、实验与结果**

此次实验，一共在七个数据集上对提出算法进行验证。得出算法的实验结果将与另外三个比较算法在相同数据集上进行对比。其中，七个数据集包括六个与论文相同的，来自UCI数据集的数据集；一个是在机器学习课程回归算法实验中用到的数据集。比较算法包括BP Neural Network，Robust AdaBoost，以及ELM。

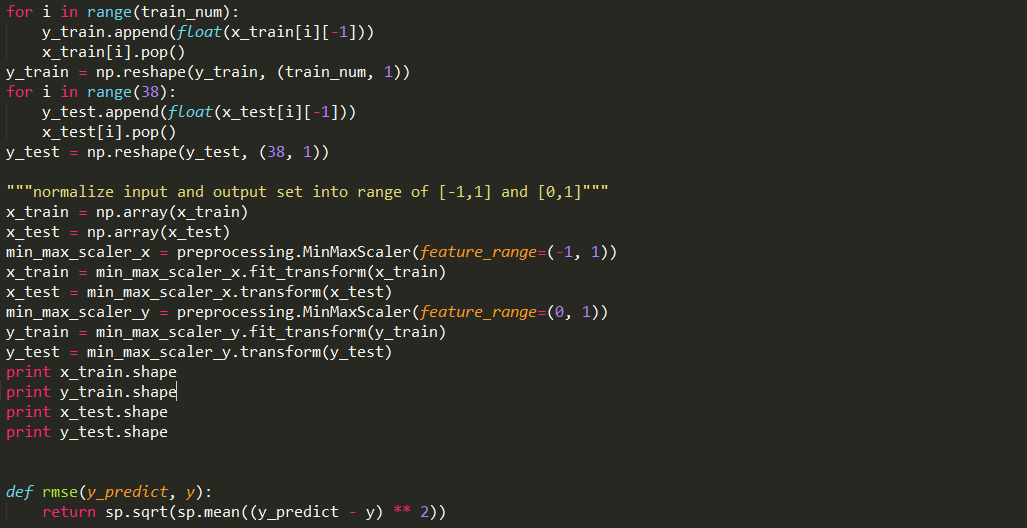
实验数据集

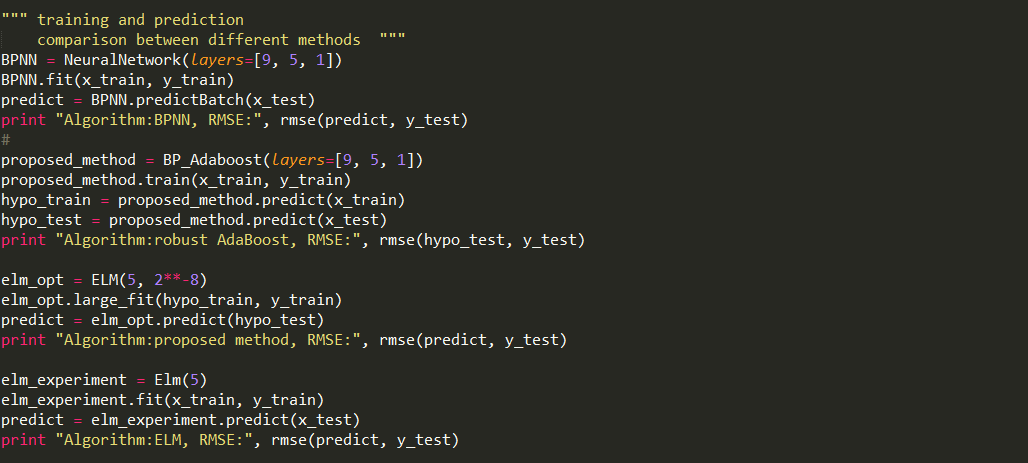
|  |  |  |  |
| --- | --- | --- | --- |
| Dataset | Training Data | Testing Data | Features |
| Cloud | 70 | 38 | 9 |
| Housing | 354 | 152 | 13 |
| Auto-MPG | 320 | 72 | 7 |
| Computer hardware | 110 | 99 | 7 |
| Breast Cancer | 114 | 80 | 32 |
| Abalone | 3050 | 1127 | 7 |
| LVST | 80 | 45 | 308 |

1. **实验主要代码截图**

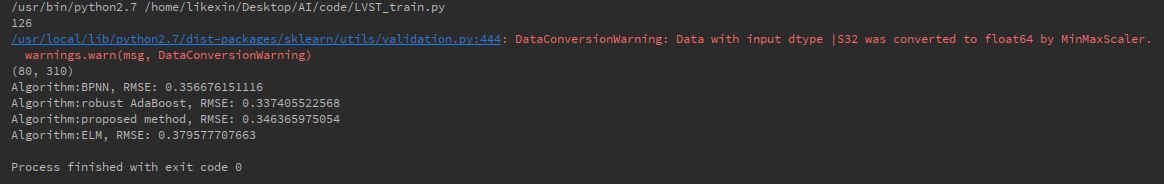
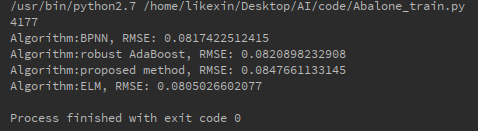
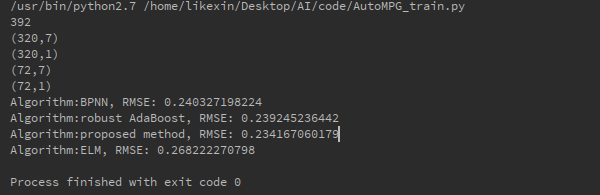
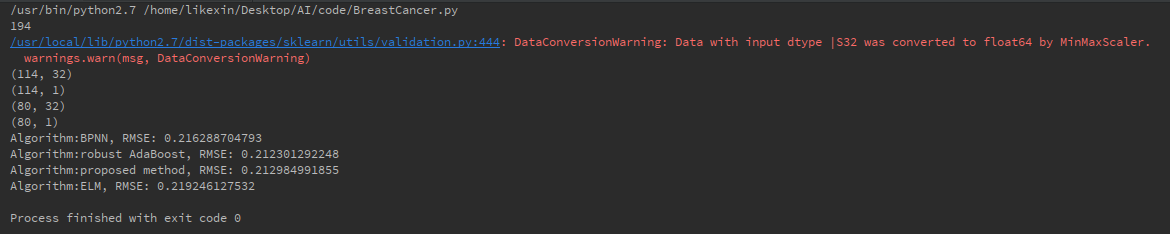
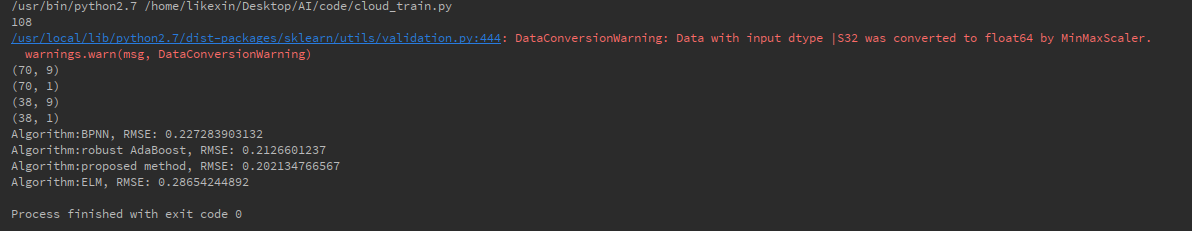
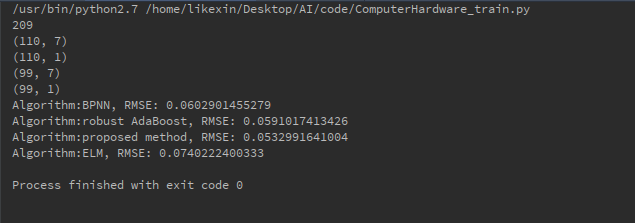
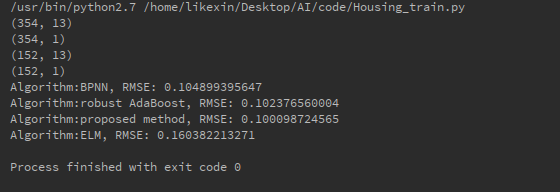
（详见cloud\_train.py，Housing\_train.py，AutoMPG\_train.py，Computer Hardware.py，BreastCancer.py，Abalone\_train.py，LVST.py）







1. **实验结果截图**

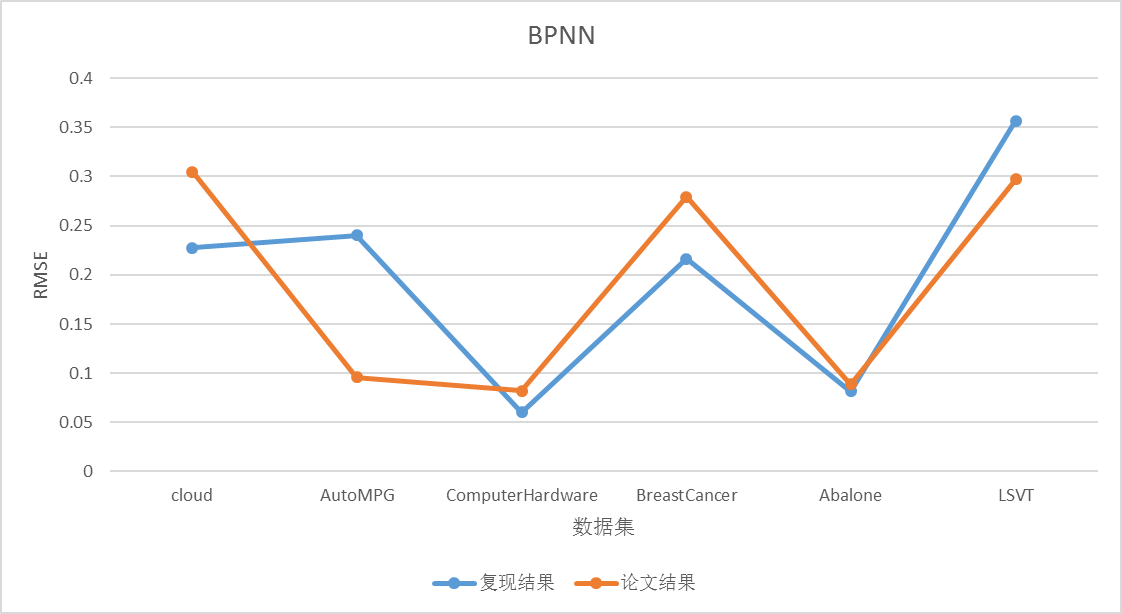
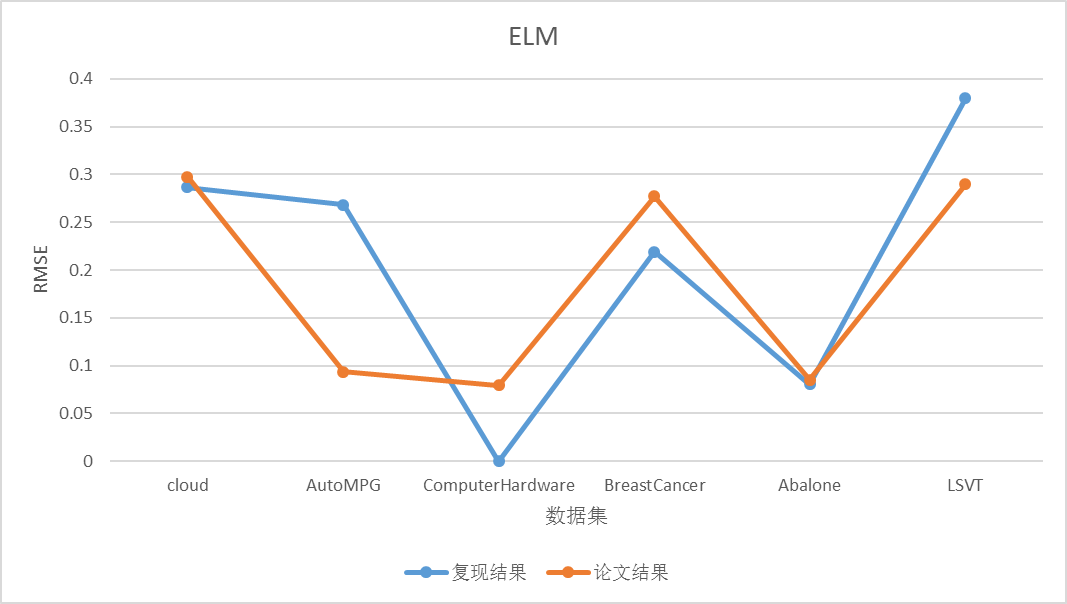
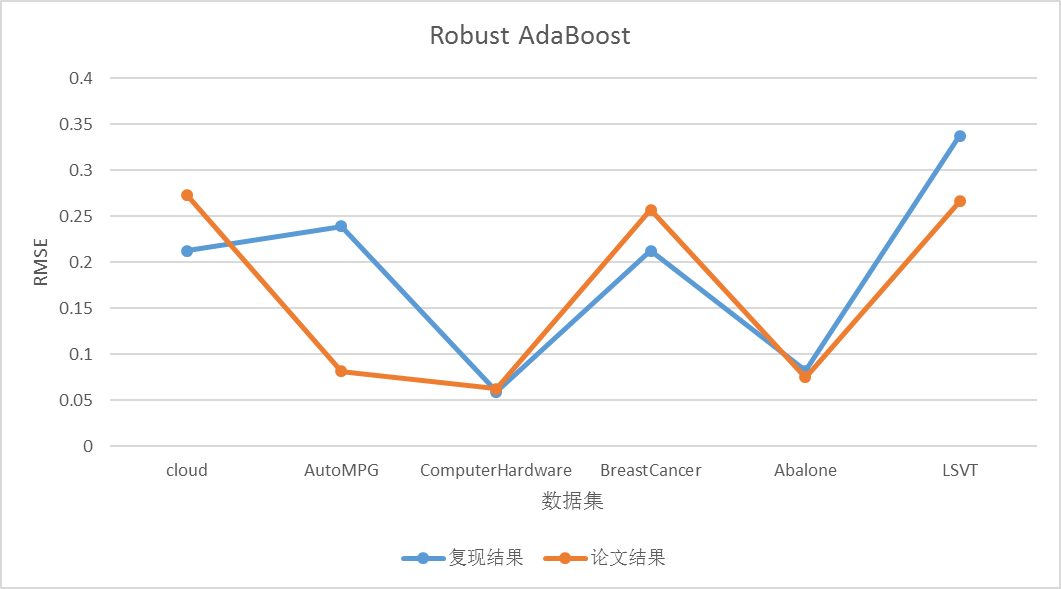
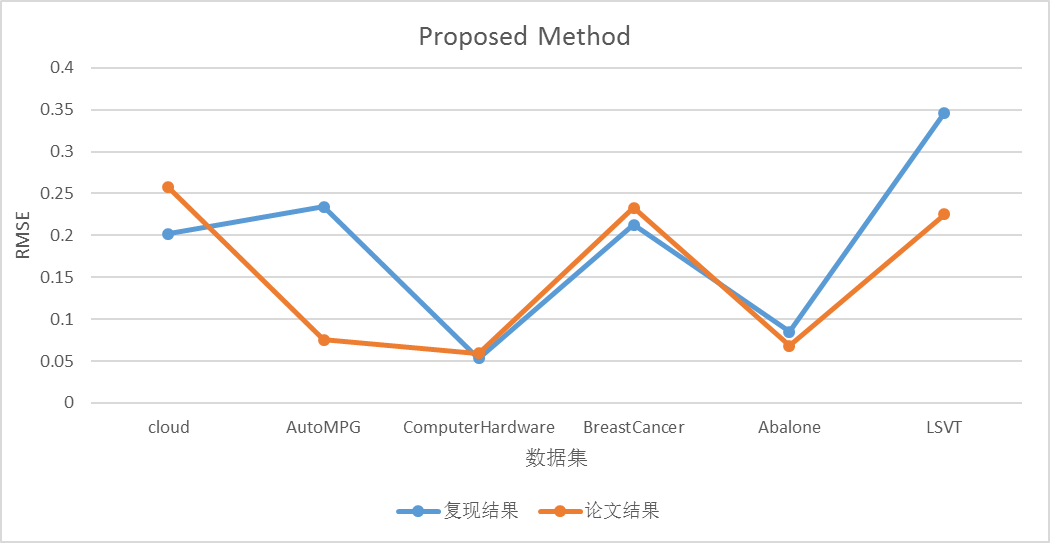


1. **衡量RMSE值的实验结果**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Proposed Method | Adaboost | ELM | BPNN |
| Cloud | 0.202134767 | 0.212660124 | 0.286542449 | 0.227283903 |
| Housing | 0.100098725 | 0.10237656 | 0.160382213 | 0.104899396 |
| Auto-MPG | 0.23416706 | 0.239245236 | 0.268222271 | 0.240327198 |
| Computer hardware | 0.053299164 | 0.059101741 | 0.07402224 | 0.060290146 |
| Breast Cancer | 0.212984992 | 0.212301292 | 0.219246128 | 0.216288705 |
| Abalone | 0.084766113 | 0.082089823 | 0.08050266 | 0.081742251 |
| LVST | 0.346365975 | 0.337405523 | 0.379577708 | 0.356676151 |

1. **复现结果与论文结果的比较**

在与论文中使用相同的数据集上，对四个方法进行结果比较。比较结果如下图所示。可以看出，复现结果与论文结果基本相近。且在67%的数据集上，复现结果有更优的表现。



1. **复现结果分析**

复现结果如下图所示。可以看出，在71%的数据集上，Proposed Method的表现都要略微优于其他三个比较算法；剩余数据集上(Breast Cancer，Abalone)，四个方法的表现几乎一致。然而，复现结果显示提出方法总体表现虽然是最优的，但它并没有带来显著的优化能力。

**五、总结**

**在机器学习的课程上接触了Adaboost算法，在实验中对于截止阈值的选择是按照实验说明的要求设定的，当时虽然对它有些困惑但也并没有对它给予过多的关注。在这次的课程任务中，通过对论文的学习，对AdaBoost阈值的确定方法有了新的认识——弱学习器在 m 个样本上的预测值的统计特点被用以自适应。**

**另外，对BP算法和ELM 算法的实现，也让我对于神经网络也有了更加具体的认识。**

**整个课程任务虽然没有想象中的困难，但从论文选择、论文研读和翻译，再到算法的理解和实现，最后进行复现实验再与论文结果相比较，整个过程也耗费了一定的时间与精力。能自己完整地完成整个过程中的所有任务，其实是非常高兴的，我也在其中体会到了很大的乐趣，对未来进行进一步的学习充满了信心和期待。**