**头脑风暴会议记录**

**1）议题：（所选题目）**

软件缺陷分析与预测系统

**2）参会人员：**

第12组：张鑫成

杨晓雪-20301055

张辰昕-20271053

张鑫成-20271055

王继航-20301019

主持人：张鑫成

记录员：张辰昕

**3）主持人发言：（解释题目）**

根据提供的软件缺陷数据集，构建软件缺陷分析与预测系统。并结合机器学习的分类算法，对不同算法性能进行比较，选出性能最好的算法。

**4）成员发表意见：（每位成员独立发表意见，每人可以发表多条意见）**

张鑫成：查阅相关论文，理解题目含义，观察提供的数据集，初步确定问题分类以及解决方案。目前讨论的重心应在于不同算法的提出上。

王继航：由于给定的数据集对于缺陷的判断是有标签的，所以本质我们的算法需要解决的是分类问题。那我们应该从我们熟悉的算法，比如KNN、SVM等入手，进行深入探讨。

张辰昕：可以先尝试上学期学过的人工智能基础课程里的一些基础算法，记录相对应的性能和效果。然后再结合本学期相对应的理论课上学到的知识，在其算法基础上进行修改，对比性能的变化。

杨晓雪：我觉得在csdn上查看一些已有的程序思路自己算法，然后我们选出几种较为优秀的算法，对算法不断进行替换，比较算法的损失函数以及一些指标。

张鑫成：尝试决策树分类法，朴素的贝叶斯分类算法、神经网络法 、KNN、以及SVM算法，在选择分类算法之后分别尝试调节损失函数以及优化函数使其达到最优值，然后初步选择出最优算法，再进一步深入，调节损失函数以及优化函数使其性能最佳。

张辰昕：先上网查阅相关论文等资料，看看其他类似项目选择了什么算法，仿照前人的经验尝试算法，记录性能结果。如果结果不理想的话再采用备选算法，对结果进行比较。

王继航：或许我们可以从我们最熟悉的几种分类算法入手，若是准确率可以达到70%以上，那再尝试通过调节相关参数使其更加高效准确；否则逐步对基本算法进行改进完善，使其正确率达到相应标准，再进行超参等修改。

**5）方案讨论排序决策过程：**

张鑫成的方案比较详实，能基本保证项目需求。

王继航的方案适合入手，但不能一定保证最后的准确率。

张辰昕的方案中论文所提及的算法可能过与复杂，不适合小组复现。

**6）结论：（讨论确定至少2个最终方案）**

方案1：尝试决策树分类法，朴素的贝叶斯分类算法、神经网络法 、KNN、以及SVM算法，在选择分类算法之后分别尝试调节损失函数以及优化函数使其达到最优值，然后初步选择出最优算法，再进一步深入，调节损失函数以及优化函数使其性能最佳。

方案2：或许我们可以从我们最熟悉的几种分类算法入手，若是准确率可以达到70%以上，那再尝试通过调节相关参数使其更加高效准确；否则逐步对基本算法进行改进完善，使其正确率达到相应标准，再进行超参等修改。