2020秋季学期 机器学习-期中大作业

**神经网络：**某型设备从投入使用到最终报废的整个寿命周期内，设备的健康状态大致可分为3个阶段：“初期磨合”、“稳定运行”和“快速老化”。为了对该设备的健康状态进行有效的监控，布置多种传感器来采集设备的运行状态。对传感器数据进行初步的信号处理，可获得一组70维的数据特征。通过模型训练学习这70维的数据特征与设备3个阶段的健康状态的关系，即可通过传感器数据的获取与分析来实现对设备所处的健康状态的监测与判断。针对上述问题，请构建如下的机器学习问题：

**数据：**，（注意作业训练数据集所给的标记为，分别对应于“初期磨合”、“稳定运行”和“快速老化”3个阶段，各小组也可以根据模型来自行设定3个阶段标记的取值。）

**模型：**多层前馈神经网络模型，各小组也可以根据资料查阅和学习后选择其他类型的神经网络模型。

作业提交模型训练代码和报告，报告回答以下问题：

（1）描述10折交叉验证对数据集的处理；（5分）

（2）描述所使用的神经网络模型；（10分）

（3）描述训练模型所使用的算法；（15分）

（4）分析模型训练结果，包括训练集错误率和测试集错误率；（10分）

（5）总结模型训练过程中的收获。（10分）

报告得分规则：写出具体的数学公式及推导过程。

代码得分规则（共50分）：不限定代码语言（若用python，请提交Jupyter notebook文件）。用测试数据集测试每个学习小组提交的模型的错误率并排序，错误率最低的学习小组代码得分50分，其余小组按排序依次减1分。请各个小组写出测试函数，与代码一起提交。

注：代码或报告雷同将严重影响最后得分。

提交时间：**2020年12月12日22:00之前**

作业格式：代码+报告（报告为pdf文件，报告首页标明小组成员姓名+学号）

作业命名（邮件名称和提交作业文件名称）：学习小组xx-机器学习-期中大作业，例如：学习小组12-机器学习-期中大作业

请各小组完成作业后，邮件形式发送至学委郑均同学的邮箱：3028208898@qq.com