暑期学校实验项目：高考志愿填报助手

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 小组名称 | AI算法A组 | | | | | | |
| 姓 名 | 高捷 | 专业 | 人工智能 | 班级 | 091181 | 学号 | 09118101 |
| 实验时间 | 2020.8.31-2020.9.23 | | 指导教师 | 孔祥龙 | | 成绩 |  |
| **一、实验背景和目的**  正值一年高考结束，志愿填报是准大学生的头等大事。本项目：高考志愿填报助手，意在帮助考生了解自己、明确目标定位，并且根据自己所在省份的预测自身适合的大学与专业。上学期，我们收集了各个省份的一分一段表以及部分高校的专业录取分数线，短学期内，我们组的任务目标为根据往年分数数据，设计神经网络模型帮助考生预测可选大学，确定其冲一冲，稳一稳，保一保的若干所大学。 | | | | | | | |
| **二、小组任务和个人任务**  小组任务：根据往年分数数据，建立神经网络模型，通过模拟数据进行训练，预测冲一冲，稳一稳，保一保的若干所学校。  个人任务：对第59-73共15所学校的数据进行模拟，生成由学生名次、学生所在省份、学生的文理类别、被录取的大学所组成的数据结果。 | | | | | | | |
| **三、个人任务需求分析**  首先根据高校各专业的分数线，按省份、类别、年份划分，确定某大学某年在某地的专业录取线最低分和最高分。然后人工地添加一个浮动分数，随机生成N个数据。N的取值可以近似于一分一段表所对应的相邻两分数人数的差值。由此生成模拟数据提供给小组其他成员进行模型训练。 | | | | | | | |
| 1. **实验过程（需附上关键代码及相关说明）**     1.预先设置好年份、省份、文理科列表。    2.数据读取。通过切片的方式取得负责的十五所高校数据，并将结果存储。    3.通过三层循环遍历数据。    4.通过sort函数对分数进行大小排列，获取高校专业录取线最高/最低分，并且让最高分加上一个浮动分数。      5.人数生成，至少一名，整除28代表在该分段取1/28学生去上某高校。    6.生成随机分数，若超出一分一段表则设为1。 | | | | | | | |
| **五、实验结果与分析**  结果如图所示，生成包含排名，省份，文理科，大学名称的数据csv文件。    这些数据是我们模拟出来的考生信息，年份信息在模拟过程中被忽略。上图显示了几名上了西南财经大学的青海理科考生的排名信息。  模拟数据环节能够完善数据集，但是由于高校对各省份录取人数等数据缺失，数据集在数量上仍有不足，因此只能近似地预测，只取1/28。如果在数据收集方面更进一步，则可选择更贴合实际的模拟人数，使得神经网络的训练结果更准确。 | | | | | | | | |
| 1. **实验总结与心得体会**   高考志愿填报系统是实用性很高的一个项目，我们组作为算法组，主要负责基于地域的高校推荐，提供给用户冲一冲，稳一稳，保一保的几所学校选择。我们选择训练一个神经网络模型来进行预测。  在组内的分工中，我负责模拟数据部分，生成由学生名次、学生所在省份、学生的文理类别、被录取的大学所组成的数据结果，为神经网络提供训练数据。  本次实验是一个较大型的多人多组合作项目，要做好这样的一个项目，不仅要代码实力过硬，更要多沟通多交流。就拿我们小组来说，在项目成立初期，大家都比较迷茫，不清楚到底要做些什么。后来我们组织了几次交流会，大家一同分析问题，各抒己见，在思想的碰撞中我们找到了前进的方向——神经网络。与此同时，各组之间的沟通也十分的重要，比如，和数据处理组的对接，和前端开发组的交流。每个小组的工作都是整个项目不可分割的一部分，只有各组之间交流通畅，才能使项目平稳运作。  当然，本次实验也存在着许多的不足。首先，在任务的分配方面，不应该把所有决策权交给单一的个人，而是应当让每个组员都参与进来，这样不仅能调动大家的积极性，更能增加大家的责任感。其次，在项目推进过程中，要敢于提出问题，提出异议。做项目时，组员间的关系就不仅仅是同学，而更是一种同事关系。只有不断地质疑，不断地为自己为组员提供适当的压力，小组合作才能顺利地进行。  由于短学期时间的有限，我们无法将这个神经网络模型训练得非常完美，它在准确性等方面仍存在一些不足，一些我们前期提出的扩展功能设想也未能成功实现。正是有了这些遗憾，我们会在接下来的编程，算法学习中投入更充足的热情。  总而言之，通过本次软件实践，我收获了许多，不仅有代码方面的知识，还有与人相处的道理。相信我能将这些学到的东西运用到接下来的学习生活中去。 | | | | | | | | |

2020年9月制