北 京 交 通 大 学

课程设计开题报告表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 课程名称 | | 大数据概论 | 学院 | 软件学院 软件工程专业 |
| 小组成员 | | 贺思超、陈杰、韩熔、解世超 | | |
| 课程设计题目 | | 基于往年气温数据预测未来天气气温 | | |
| 课程设计研究内容及  选题依据 | 天气预测一直以来就是人们的愿景以及痛点之一，有效的天气预测有利于政府对资源的统筹调度。  本实验拟采用Python语言，从网页爬取气温数据，使用Pandas和NumPy对气温数据进行处理与分析，并采用Sklearn框架进行机器学习处理。最后使用TensorFlow框架构建ANN人工神经网络进行建模预测。 | | | |
| 课程设计关键问题预期目标 | 基于天气预测的广泛应用和重要意义，以及其与大数据极强的关联性，本小组选择气温预测进行课程研究，旨在运用大数据分析与挖掘的理论与技术解决气温预测的相关问题。 | | | |
| 课程设计的研究计划及预期解决方案 | 要包括数据来源、存储（excel或是数据库都可以）、模型、评估、可视化等。以上内容是小组内部讨论过的，是暂时的方案，后续可以更改，我需要了解你们的思路是否正确。 | | | |
| 技术难点及  解决方案 | 技术难度在于如何减小预测的误差，训练批次的增加和神经元个数的增加可能会导致过拟合问题。  解决方案有1.增加训练集；2.减小神经元个数，降低复杂度；3.正则化 | | | |
| 组内分工及  开题  当前进展 |  | | | |