课程预习报告目录

1 预习目标 (对应课程目标1)

1.1对这门课的期望

我期望在机器学习课程中，能够深入了解机器学习的基本概念和算法，例如监督学习、无监督学习、强化学习等。我希望能够详细了解各种机器学习算法的工作原理、何时以及为何使用它们。此外，我还期望学习更多的机器学习方法，例如集成学习、特征选择、降维、深度学习等。同时我期望能获得更多的项目实践经验，提高自己的代码能力，为将来科研或工作中应用机器学习打下基础。

1.2对人工智能的理解

在上课之前，我对人工智能的理解主要是基于其在电影和小说中的形象，以及一些日常生活中的应用，如语音助手、自动翻译、推荐系统等。我理解人工智能是一种能够模仿人类智能的技术，可以像人类一样学习和适应环境，甚至在某些方面超越人类的能力。

我对人工智能的期望是希望能够看到它在更多领域得到广泛应用，并为人类社会带来更多的便利和进步。我希望人工智能够像人类一样更加智能，能够更好地理解和处理复杂的任务，如自然语言处理、推理、创造等。同时，我也关注人工智能的伦理和社会影响，希望它能够被合理、公正地使用，并避免产生负面影响。我相信随着技术的不断进步和应用场景的不断扩展，人工智能将会在更多领域发挥重要的作用，并为人类社会带来更多的变革和发展。

1.3预习计划

1.安装anaconda，配置pytorch环境并安装pytorch库，并在该环境下安装如pandas、numpy、scikit-learn等需要用到的库

2、跟随mooc上的python语言程序设计课程进行编程基础学习

3、之后从Kaggle网站上下载泰坦尼克号数据集，跟随教程敲一遍代码，将每一步都弄清楚，获得自己的第一个预测模型，了解最基本的数据处理方法和模型构建方法（第三方库的调用）

4、如果有精力则更深入的了解机器学习的算法

2 基础知识预习 (对应课程目标6、7和8)

在基础知识预习过程中，我学会了如何设置和使用Python环境，如何在命令行中运行Python代码，以及如何使用Python中的基本语法和数据类型。我还学会了如何定义和使用函数，以及如何使用控制结构来控制程序的流程。此外，我也学会了如何使用Python的模块和库来执行各种任务（学会了如何处理文件等），学会了如何定义和创建类和对象。最重要的，是了解了自顶向下的程序设计方法和面向对象的编程思想

3 预习项目设计(对应课程目标2，3)

3.1 项目选题

1.选题：泰坦尼克号幸存者预测

2.选题原因：是最基础的预测模型，对于算法要求不高，是本课程的第一个任务，网络上有丰富的开源资源方便学习，能帮助我建立一个直观的完整的数据清理、模型构建过程的逻辑架构，并且在kaggle上参加竞赛形式有助于增强学习动力

3.2 程序设计

1.**读取数据：**首先需要获取泰坦尼克号乘客的数据，包括他们的年龄、性别、舱位等级和其他可能影响他们是否幸存的因素。

2.**数据清洗：**接着，需要清洗数据，移除或修正错误、遗漏或异常值。这可能包括处理缺失值、异常值和不符合预期的输入。

3.**数据分析：**然后，可以通过数据分析技术来研究这些数据，以找出影响幸存率的因素。这可能包括统计分析和机器学习方法。

4.**模型构建：**基于数据分析结果，可以调用sklearn库中的算法构建一个预测模型，该模型可以根据乘客的特征预测他们幸存的可能性。

5.**预测结果：**需要评估模型的准确性。这可以通过比较模型预测的结果和实际的结果来完成。

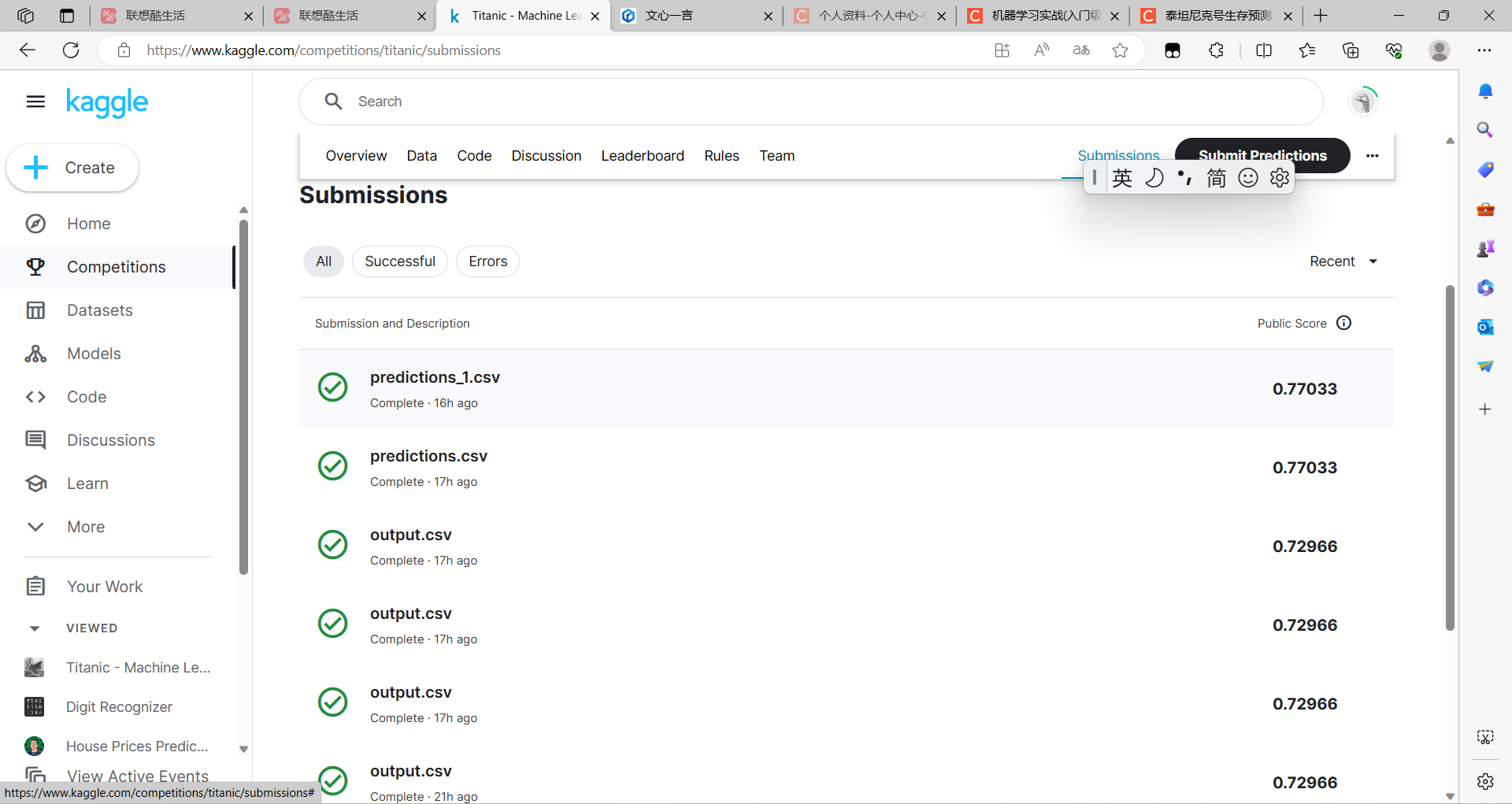
6.**模型优化：**如果模型的准确性不够高，我需要优化我的模型。这可能包括调整模型参数、添加新的特征或使用不同的算法。

3.3 运行结果

得到一个csv.文件包含一行标签和418行预测结果，提交到kaggle获取正确率，得到的正确率偏低，仍待进行进一步优化。

存在问题：1）因为对算法缺少了解，不太会调参数，导致程序警告达到最大迭代次数，没有收敛到最优解

2）pandas库中的append方法警告



4 预习总结 (对应课程目标3)

在基础知识预习过程中，我学会了如何设置和使用Python环境，如何在命令行中运行Python代码，以及如何使用Python中的基本语法和数据类型。我还学会了如何定义和使用函数，以及如何使用控制结构来控制程序的流程。此外，我也学会了如何使用Python的模块和库来执行各种任务（学会了如何处理文件等），学会了如何定义和创建类和对象。最重要的，是了解了自顶向下的程序设计方法和面向对象的编程思想。

在泰坦尼克号幸存者预测项目学习中，我获得了对机器学习程序编写过程的直观认识和了解，对应用python及其第三方库的强大功能充满憧憬。

附件2 课程预习报告评分表

|  |  |  |  |
| --- | --- | --- | --- |
| **过程要素** | **评分依据** | **分值** | **得分** |
| 1、对课程的过程和收获期望清晰、合理，选课动机明确 (6分) | （1）能够明确的给出对课程学习过程和学习结果的期望，同时期望合理，符合课程主题、符合课程学时，与预习作业要求匹配，选课动机明确 | 6 |  |
| （2）能够给出对课程学习过程和学习结果的期望，同时期望较为合理，大概符合课程主题或课程学时或与预习作业要求匹配，选课主动 | 4 |
| （3）给出的期望不太合理，选课动机不明确 | 2 |
| （4）期望完全无法匹配课程，随机或被动选课 | 0 |
| 2、预习计划明确，工作量充足，对预习作业中的要求能够满足，能够支撑掌握Python的基本使用技巧 (6分) | （1）预习计划明确，工作量充足，对预习作业中的要求能够满足，包含了基本的Python工具使用和语言使用，有最终作品计划 | 6 |  |
| （2）计划一般明确，工作量比较充足，包含工具和语言使用，有最终作品计划 | 4 |
| （3）计划随意，工作量不足，工具语言有但不明确，没有最终作品计划 | 2 |
| （4）没有表述 | 0 |
| 3、Python基础知识掌握(6分) | （1）对关键的python基本概念有正确的描述，关键知识点较为全面 | 6 |  |
| （2）对关键的python基本概念有基本正确的描述，关键知识点不全面 | 4 |
| （3）知识点描述不正确，基本不包含关键基本概念 | 2 |
| （4）没有表述 | 0 |
| 4、预习项目设计合理(6分) | （1）预习项目工作量饱满，技术含量高，选题合理，充分运用了预习的知识点 | 6 |  |
| （2）预习项目工作量较为饱满，有一定技术含量，选题一般，运用了预习的知识点 | 4 |
| （3）预习项目工作量不饱满，基本没有技术含量，选题一般，运用预习知识点很少 | 2 |
| （4）没有表述 | 0 |
| 5、对预习过程有中学到的知识有明确的认识(6分) | （1）明确知道自己已经学到了哪些内容，可以完成哪些任务，对未来学习有明确计划和期望目标 | 6 |  |
| （2）对自己学到的知识认识不准确，但是对未来学习有明确的计划和期望 | 4 |
| （3）对自己学到的知识认识不准确，对未来学习有计划和期望，但不太合理 | 2 |
| （4）没有表述 | 0 |
| 总分 |  |  |  |
| 评 语  批阅签名： | | | |