《机器学习》第11周实验课流程

2022年4月26日

一、Openbayes 实验环境搭建

链接: https://openbayes.com/docs/getting-started/

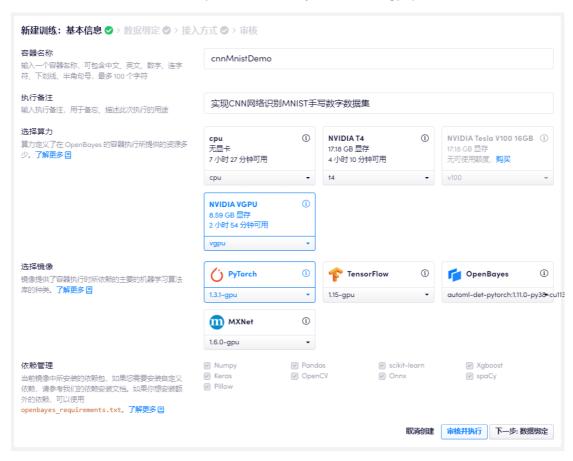
1. 注册 openbayes

网址 https://openbayes.com

2. 在页面左侧选择模型训练



3. 在界面中选择如图所示环境(NVIDIA VGPU, PyTorch-1.3.0-gpu)



4. 在弹出的页面中点击下一步



5. 点击下一步



6. 点击执行



7. 在工作空间界面中导入实验的文件 char_rnn_classification_tutorial.ipynb, data.zip

二、姓名分类

1. 任务描述:

CLASSIFYING NAMES WITH A CHARACTER-LEVEL RNN:

构建和训练一个基本的 RNN 来对单词进行分类。

将单词读取为一系列字符,在每一步输出预测和"隐藏状态",将其先前的隐藏状态馈送 到每个下一步。我们将最终预测作为输出,即单词属于哪个类。

数据集包含来自 18 种语言的几千个姓氏, rnn 模型训练完成后, 根据拼写预测一个名字来自哪种语言

> Dovesky

(-0.20) Russian

(-2.82) Polish

(-2.96) Czech

> Jackson

(-0.78) Scottish

(-1.63) Russian

(-1.66) English

> Satoshi

(-1.22) Japanese

(-1.53) Italian

(-2.38) Polish

2. 流程:

- (1) 数据读入、预处理
- (2) 初始化模型、优化器
- (3) 训练、测试
- 3. 参考: https://pytorch.org/tutorials/intermediate/char_rnn_classification_tutorial.html https://www.bilibili.com/video/BV1Ui4y1t78N?p=1
- 4. 代码: char_rnn_classification_tutorial.ipynb
- 5. 数据集: https://download.pytorch.org/tutorial/data.zip

三、作业安排

完成上面姓名分类任务, 5月10日之前厦门大学SPOC平台中提交一份实验报告。