

# 斯坦福CS231n课程作业# 1简介



杜客 8 个月前

译者注：本文[智能单元](#)首发，由[杜客](#)翻译自斯坦福CS231n课程作业1介绍页面[\[Assignment #1\]](#)。该课程共有3个作业，建议深度学习入门的知友完成。

## 原文如下

在本作业中，将基于k-最近邻（k-Nearest Neighbor）或者SVM/Softmax分类器实践一个简单的图像分类流程。本作业的目标如下：

- 理解基本的**图像分类流程**和数据驱动方法（训练与预测阶段）。
- 理解训练、验证、测试分块，学会使用验证数据来进行**超参数调优**。
- 熟悉使用numpy来编写向量化代码。
- 实现并应用k-最近邻（**k-NN**）分类器。
- 实现并应用支持向量机（**SVM**）分类器。
- 实现并应用**Softmax**分类器。
- 实现并应用一个**两层神经网络**分类器。
- 理解以上分类器的差异和权衡之处。
- 基本理解使用**更高层次表达**相较于使用原始图像像素对算法性能的提升（例如：色彩直方图和梯度直方图HOG）。

## 安装

有两种方法来完成作业：在本地使用自己的机器，或者使用[Terminal.com](#)的虚拟机。

## 云端作业

Terminal公司为我们的课程创建了一个单独的子域名：[www.stanfordterminalcloud.com](#)。在该域名下注册。作业1的快照可以在[这里](#)找到。如果你注册到了本课程，就可以联系上助教（更多信息请上Piazza）来得到用来做作业的点数。一旦你启动了快照，所有的环境都是为你配置好的，马上就可以开始作业。我们在Terminal上写了一个简明[教程](#)。

## 本地作业

点击[此处](#)下载代码压缩文件。初次之外还有些库间依赖的配置：

**[选项1]使用Anaconda：**推荐方法是安装[Anaconda](#)，它是Python的一个发布版，包含了最流行的科研、数学、工程和数据分析Python包。一旦安装了它，下面的提示就都可略过，准备直接开始写作业吧。**译者注：推荐。**

**[选项2]手动安装，虚拟环境：**如果你不想用Anaconda，想要走一个充满风险的手动安装路径，那么可能就要为项目创建一个[虚拟环境](#)了。如果你不想用虚拟环境，那么你的确保所有代码需要的依赖关系都是景在你的机器上被安装了。要建立虚拟环境，运行下面代码：

```
cd assignment1
sudo pip install virtualenv          # This may already be installed
virtualenv .env                      # Create a virtual environment
source .env/bin/activate             # Activate the virtual environment
pip install -r requirements.txt      # Install dependencies
# Work on the assignment for a while ...
deactivate                           # Exit the virtual environment
```

**下载数据：**一旦得到作业初始代码，你就需要下载CIFAR-10数据集，然后在**assignment1**目录下运行下面代码：**译者注：也可手动下载解压后放到cs231n/datasets目录。**

```
cd cs231n/datasets
./get_datasets.sh
```

**启用IPython：**得到了CIFAR-10数据集之后，你应该在作业**assignment1**目录中启用IPython notebook的服务器，如果对IPython notebook不熟悉，可以阅读[教程](#)。

**注意：**如果你是在OSX上的虚拟环境中工作，可能会遇到一个由matplotlib导致的错误，原因在[这里](#)。你可以通过在**assignment1**目录中运行start\_ipython\_osx.sh脚本来解决问题。

## 提交作业

无论你是在云终端还是在本地完成作业，一旦完成作业，就运行collectSubmission.sh脚本；这样将会产生一个assignment1.zip的文件，然后将这个文件上传到你的dropbox中这门课的[作业页面](#)。

## Q1：k-最近邻分类器（20分）

IPython Notebook文件knn.ipynb将会带你实现kNN分类器。

## Q2: 训练一个SVM (25分)

IPython Notebook文件`svm.ipynb`将带你实现SVM分类器。

## Q3: 实现Softmax分类器 (20分)

IPython Notebook文件`softmax.ipynb`将带你实现softmax分类器。

## Q4: 实现2层神经网络 (25分)

IPython Notebook文件`two_layer_net.ipynb`带你实现一个2层神经网络。

## Q5: 更高层次表达: 图像特征 (10分)

IPython Notebook文件`features.ipynb`带你比较使用更高层次表达相较于使用原始像素对于算法性能的提升。

## Q6: 加分: 做点儿其他的! (+10分)

实现、调查或者分析其他一些与本次作业相关的主题, 并使用你实现的代码。例如, 有没有什么你们可以问的有趣问题? 能不能做出一些具有洞察力的图表? 或者任何有趣且值得一看的东西? 也许你还可以对损失函数做点其他实验? 如果你尝试了一些够酷东西, 我们将给你10分的加分, 这将影响你的课程表现。

全文完。

## 译者反馈:

1. 转载须全文转载并注明原文链接, 否则保留维权权利;
2. 如对翻译有意见建议, 请通过评论批评指正, 贡献者均会补充提及;
3. 知行合一, 建议深度学习入门阶段的知友用心实现作业, 将有收获;
4. 后续我将根据作业内容和自己的学习笔记原创教程。