

# 面向“深度学习”课程的 Python 实践型教学探索

金国哲 崔旭

(延边大学工学院智能信息处理研究室 吉林延吉 133002)

**【摘要】** 本文探索一种基于 python 项目驱动的深度学习教学模式,并应用于高年级本科生及研究生教学过程中,实践证明这种模式可以加深学生对深度学习中基本理论的理解,同时可以激发学生的研究兴趣,为更深层次研究工作打好基础,提高教学效果。

**【关键词】** python; 项目驱动; 深度学习

**DOI:** 88888888888888888888888888888888

2018年4月,为了引导高等学校瞄准世界科技前沿,不断提高人工智能领域科技创新、人才培养和国际合作交流等能力,为我国新一代人工智能发展提供战略支撑,教育部特制定了《高等学校人工智能创新行动计划》。本文围绕人工智能领域人才培养体系建设的问题,结合延边大学计算机教育的特点,探索实践型“深度学习”课程教学改革的问题。

## 1、教学目标

针对高年级本科生及研究生,围绕深度学习中梯度下降、优化算法、神经网络结构设计等核心问题,通过项目驱动的方式,使学生掌握通过 python 处理大规模数据、搭建深度学习模型、调试模型的方法。实践教学的先决条件是学生需要掌握基本的 python 语言,机器学习理论。

Python 语言因语法简单、容易上手著称,是用户数增加最快的一种编程语言。这与近年人工智能的发展,特别是深度学习技术的发展密不可分。深度学习的概念是由 Hinton 等人于 2006 年提出<sup>[1]</sup>,是机器学习领域的一个分支。深度学习网络通常是含有多层感知机的神经网络结构,通过深度网络可以自动抽取底层特征,并通过组合形成更为抽象的高层特征,可用于分类、回归任务。伴随着高性能 GPU 计算条件的成熟和互联网提供的海量数据,深度学习技术近些年得到了迅猛的发展,并逐渐在工业界得到应用。例如谷歌的机器翻译系统就是通过深度学习中的 seq2seq 模型<sup>[2]</sup>训练得到的。

Python 为深度学习提供了良好的开发生态和平台支撑,Theano、Tensorflow、Pytorch 等主流的深度学习框架支持的主要语言接口均是 Python。根据计算图的组织方式的不同,深度学习框架可分为静态计算图框架和动态计算图框架。Theano 和 Tensorflow 均采用了静态计算图,而 pytorch 采用的是动态计算图。相比于静态计算图,动态计算图更直观、便于调试、学习曲线平缓等优点,因此,结合学生群体的特点,本教学实践中采用了 pytorch 框架。

## 2、深度学习课程体系的设计

### 2.1 课程主要讲授内容

本次教学重点是在培养学生对深度学习的基本概念的理解及应用基本概念解决实际问题的能力。整个课程体系分为基础和进阶两部分共 7 个项目构成,其中基础部分由线性回归、罗杰斯特回归、softmax 分类器构成。进阶部分由多层神经网络 MNIST 分类、卷积神经网络<sup>[3]</sup>做 MNIST 分类、LSTM 网络做情感分析<sup>[4]</sup>,基于 seq2seq 模型的机器翻译构成。

### 2.2 基础部分

首先在基础部分培养学生重点掌握深度学习的基础知识和 pytorch 基本操作。在课程前期阶段需要以线性回归、罗杰斯特回归,softmax 分类器等基本模型作为切入点,使学生掌握

利用 pytorch 构建深度学习模型的基本步骤。步骤如下:

Step1: 对数据预处理,生成数据加载类:真实场景下,根据深度学习应用场景不同,数据有图片、音频、视频、文本等多种形式,数据的格式也是千差万别。以文本为例,数据可以是 txt、csv、xml 格式的文件存储。通过 python 预处理数据的一般流程如下。

- 1) 读入文本数据,构建词典,如果是中文还需要进行分词。
- 2) 根据词典将数据索引化,用词的整型索引集合表示句子。
- 3) 将文本数据按照一定的比例分割成训练集、验证集、测试集。
- 4) 构建代表数据集的类,用于训练过程中随机打乱顺序,同时按照批次返回数据,该过程可以借助 pytorch 的 DataLoader 完成。

根据具体任务的不同,上述过程需要做调整,例如情感分析中标签部分需要按照类别个数索引化,而机器翻译中不存在标签,模型输出同样是一条句子,因此一个数据实例由一对句子构成。

Step2: 通过 pytorch 构建深度学习模型,主要是通过代码描述模型的结构。Pytorch 提供了一些基本模块,例如多层前馈神经网络中可以使用 pytorch 内建的 Linear 模块,循环神经网络则可以使用 LSTM 模块等。

Step3: 构建模型的代价函数,并配置所使用的优化算法。其中常用的代价函数,例如均方差代价函数、交叉熵代价函数等均可以直接使用 pytorch 的实现。

Step4: 构建模型的训练过程:需要借助 DataLoader,将批量化的数据提供给模型,用于训练一个步骤,再利用 pytorch 提供的自动求导机制,将梯度从输出传回至输入,之后更新模型的参数。

基础部分的教学过程中不用刻意强调数据,所有实验均可采用一些生成的模拟数据,例如线性回归模型中,可以手动规定一条直线,并通过 python 的随机算法生成该直线周围一定方差内的样本点数据。

### 2.3 进阶部分

首先,数据均采用真实存在的公开数据集,例如多层神经网络环节开始引入经典的手写字识别公开数据集 MNIST。这部分开始重点强调不同的模型在处理不同数据集或问题时表现出的差异。例如,在讲解多层神经网络过程中,可构建多种宽度和深度的隐层,用于深入分析模型复杂度、参数个数对模型性能的影响,同时自然地引入模型过拟合的概念,使得学生对深度神经网络有直观的认识,为以后的模型设计,调整参数打好基础。

其次,利用多层神经网络、循环神经网络、卷积神经网络分别建模 MNIST 数据集,来预测未知(测试数据集)手写字,并计算预测的准确率。通过这些例子,重点讲解不同结构的模型的性能差异,让学生认识到深度学习模型改进方面的必要性,同时激发学生对深度学习研究的兴趣。

再次,介绍情感分析问题及真实应用场景,并讲解利用循环神经网络做句子建模的过程,特别是需要重点说明双向LSTM网络在句子建模中的经典应用。LSTM实现过程中需要注意长度问题,像MNIST数据集中每张图片的大小都是固定的,因此可以采用固定长度的LSTM网络,而像情感分析问题中每个样本实例的句子长度都是可变的,因此需要动态LSTM完成句子建模。相比于tensorflow,pytorch的动态LSTM稍显麻烦一些,需要一个批次的句子按照句子长度的降序提前排好序,再做填充,输入到模型,从模型输出后还需要恢复成原始的形式。如果学生希望做自然语言方面的进一步研究,动态LSTM的实现是必不可少的环节,需要熟练掌握。

最后,介绍seq2seq模型的Encoder和Decoder结构,同时介绍Encoder和Decoder的几个变种,例如Encoder的LSTM层叠形式,双向LSTM结构,Decoder过程中增加针对Encoder输出的注意力模型[5]等,并提供学生扩展阅读的论文链接。以基本的英法翻译语料库为例,通过pytorch实现基于seq2seq模型的翻译系统,使学生了解当今最先进的翻译系统的基本原理,对机器翻译有基本的认识。重点讲解Decoder和Encoder之间的联系、Decoder的实现过程中采用的teacher forcing机制等实现细节。

#### 2.4 课程讲授方式及实践环节

课堂讲授方式上,所有上述内容均采用python的jupyter交互式文档工具。jupyter将深度学习理论和代码讲解部分用文本形式展现,同时在文档中嵌入需要讲解的代码,每一步代码均可独立运行,另外结合pytorch提供的动态计算图特性,jupyter可立即展现代码运行的结果。这些特性使学生第一时间对理论讲解有直观的认识,有利于学生课后复习和实践。

学生实践环节上,结合延边大学少数民族特色,设计实践环节项目具体如下:(1)利用LSTM构建朝鲜语分写预测模型,权重为0.2;(2)利用seq2seq构建朝鲜语形态素分析和词性标注模型,权重为0.3;(3)基于CNN的朝鲜语情感分析,权重为0.3;(4)Cifar10图片分类任务,可用任意模型实现,权重为0.2。

实践环节中为了提高学生的参与度和质量,将班级学生按3人一组分组,每组成员需要分工,并在报告书中明确说明每个学生的贡献。评价标准按照项目的完成度和最后的项目报告书质量给出。项目报告书属于主观评价,项目完成度根据模型具体的表现量化,例如每个题目都会给出一个基准模型的性能评估指标,实际的分数与这个评估指标做对比,用分段的方式给出。

### 3、结语

当前市场对人工智能人才需求巨大,为了兼顾工程型和研究型深度学习人才的培养需求,结合延边大学的少数民族大学特色,建立了一种基于python的实践型深度学习课程体系。针对高年级本科生和研究生的为期2个学期的教学实践表明,本文采用的教学体系适合于掌握一定人工智能及编程基础的学生,可以激发这些学生的人工智能相关知识的学习积极性,同时,通过课程为学生的科研活动及工程实践打好了良好的基础。

**项目信息:** 延边大学校企合作项目(延大科合字2019[34]号)

### 参考文献

- [1] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. science, 2006, 313(5786): 504-507.
- [2] Sutskever I, Vinyals O, Le Q V. Sequence to sequence learning with neural networks[C]//Advances in neural information processing systems. 2014: 3104-3112.
- [3] LeCun Y, Bengio Y. Convolutional networks for images, speech, and time series[J]. The handbook of brain theory and neural networks, 1995, 3361(10): 1995.
- [4] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural computation, 1997, 9(8): 1735-1780.
- [5] Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate[J]. arXiv preprint arXiv:1409.0473, 2014.