

# 机器学习实验课教学设计研究\*

秦婧

沈阳医学院医学信息工程学院, 沈阳 110034

**摘要** 针对机器学习实验课设计中缺乏实验之间的衔接、教学形式单一等教学效果不理想的问题, 设计基于通用实验框架的案例式实验课以及综合考核标准。在实验课的教学设计中, 基于机器学习理论课的授课内容设计4次实验课, 实验案例全部在通用实验框架中实现, 由学生自主调整参数优化模型并完成综合实验报告。通过考核结果表明, 机器学习实验课的教学设计在有效地巩固学生对理论知识理解的同时, 提高了学生对机器学习模型的学习兴趣。

**关键字** 机器学习, 教学设计, 实验教学, 通用实验框架

## A Research on Teaching Design of Experiment in Machine Learning Course

Jing Qin

College of Medical Information Engineering of Shenyang Medical College,  
Shenyang 110034, China;  
50485977@qq.com

**Abstract**—In response to the shortcomings of experiment in machine learning course characterized by experimental incoherence and a monotonous teaching approach, we have introduced a case-based experimental course structured around a universal framework. Coupled with a comprehensive assessment criterion, this instructional design aligns with the theoretical machine learning course and encompasses four integrated experiment sessions. Students are motivated to autonomously fine-tune parameters, enhance model performance, and produce exhaustive experimental reports. The assessment outcomes demonstrate that this pedagogical strategy not only solidifies students' grasp of theoretical principles but also significantly amplifies their enthusiasm for engaging with machine learning models.

**Keywords**—Machine Learning, Teaching Design, Experiment Teaching Universal Framework

### 1 引言

在人工智能、大模型时代, 高校中人工智能方向的课程已经成为学生选修的热门课程。机器学习课程作为人工智能方向的敲门砖, 不仅是大多数高校中计算机相关专业的必修课, 也是非计算机专业(例如, 财经类专业, 医学类专业)的选修课。近年来, 随着机器学习课程教学方法的改革, 从最初的偏重理论教学, 逐步发展到理论和实践相结合的教学方法, 并将实践部分的内容作为侧重点。目前, 大多数机器学习课程的授课形式均为理论与实践相结合的形式, 实践环节是对学生能否掌握机器学习理论的重要环节。比如, 叶景贞等人<sup>[1]</sup>针对本科生的机器学习课程实践, 提出了让学生采用不同语言和不同平台的灵活性的实践方式, 同时鼓励学生以小组的形式完成实验项目, 从而在学生的动手能力基础上提高了学生的参与度与团队合作精神。李阳<sup>[2]</sup>针对企业对机器学习课程考察的

要求, 结合高职学生的特点, 由浅入深地引入与实际应用相结合的案例, 激发学生的学习兴趣。王敏<sup>[3]</sup>等人针对机器学习课程中的朴素贝叶斯方法设计了教学案例, 在案例中利用爬虫来获取数据集, 并使用朴素贝叶斯方法对其分类, 实践表明该案例简单易懂, 适用于机器学习的实验课。徐国艳<sup>[4]</sup>等人在机器学习课程的实践中采用了百度的AI Studio实验平台, 学生在该平台上完成机器学习的课程实践。黄美燕<sup>[5]</sup>针对生物医学工程专业的学生, 在机器学习的实验中引入了医学案例, 有利于让学生了解机器学习在其所学的专业领域的实际应用。

综上, 针对不同专业、不同学习层次的同学, 机器学习实验教学均可以有效提高学生对理论知识的理解, 并能让他们了解机器学习模型在实际环境中的应用。基于此, 本文将机器学习实验课作为机器学习课程中的重要环节, 通过分析机器学习实验课的教学现状, 设计了基于通用实验框架的案例式的机器学习实验课以及综合考核标准。

\* **基金资助:** 本文得到全国高等院校计算机基础教育研究会项目资助(2024-AFCEC-542)。

## 2 机器学习实验课的教学现状

机器学习课程是一门理论性和实用性都很强的课程，是未来从事人工智能领域研究的基础课程。针对机器学习课程实用性强的特点，机器学习实验课设计也受到了重点关注。在机器学习课程的教学大纲中，学时安排专门设置了机器学习实验课（课内实验）的学时。由于机器学习模型的训练过程以及参数调优的时间过长，无法在给定的课内实验时间完成，在授课内容的安排中让学生利用课外时间以小组的形式完成案例学习，以此来补充课内实验时间。机器学习实验课开设以来，在授课教师和学生的授课效果反馈结果中，主要存在以下 3 方面问题。

### 2.1 教材的选择缺乏针对性

机器学习的理论课与实验课的授课教师通常是同一位教师，在安排实验课内容时，其主要是依据理论课的知识点布置实验内容，大部分的实验内容均来自理论课的教材。在实验课教学设计中，授课教师通常不会重新选择一本实验教材或者重新设定实验任务。基于此，机器学习实验课上的实验内容很难激发学生的学习兴趣，继而导致学生无法将理论课所学习的内容应用到实际应用中。

### 2.2 实验课的内容缺乏衔接

在机器学习实验课的案例设计中，大部分实验内容的设计都是基于理论课的知识点来设计相应的实验课内容，每个实验的内容相对独立，即每次实验课，仅考虑使用一个模型完成分类或回归的任务。此外，在每次实验中所用到的数据集也完全不同，未考虑到每个实验中不同模型在相同数据集上的预测效果。这样的实验课设计方式导致了学生对实验中所用到的理论知识理解不深入，无法形成完整的知识体系，影响学习效果。

### 2.3 教学形式缺乏多样性

实验课的教学形式通常采取的是讲练结合的形式。首先，由授课教师给出实验任务以及模型构建的示例代码，然后，由学生根据授课教师给出模型构建的示例代码完成实验任务。机器学习实验课的目标是让学生能够熟练应用模型，并熟知如何调整并优化参数，但通过这种传统的实验教学形式，学生虽然能完成实验任务，但缺乏对模型应用的进一步思考和探索，影响学生对模型的深入理解。

## 3 考核标准与课时安排

机器学习课程主要讲授的是机器学习的基础理论、模型以及实现方法，内容包括机器学习绪论、决策树、

模型评估、线性模型、支持向量机以及神经网络。在机器学习课程教学大纲中，课程总学时为 32 学时，考试形式为考查，其中理论课是 24 学时，实验课是 8 学时。为了让学生更加重视实验课，将考核标准中实验课成绩所占比例设置为 80%，即机器学习课程的总成绩是平时成绩与实验成绩构成，其中，平时成绩占 20%，实验成绩占 80%。由此可见，实验课是机器学习课程中的一个重要组成部分，实验课的目标是将理论课的知识应用到真实数据集中实现预测或分类的任务。实验成绩由实验代码和综合实验报告两部分构成，分值比例各占 50%，综合实验报告是基于 4 次实验结果的对比分析包括模型的优缺点、结果的可视化分析、实验中所遇到的问题汇总。具体的考核标准见表 1。

表 1 机器学习实验课的考核标准

考核形式	考核大项	子项（权重）	总分值
实验考核	4 次实验课的 实验代码	按时提交（10%） 模型的代码实现（40%） 预测结果评价（10%） 预测结果的可视化（20%） 模型调优（20%）	40
	综合实验报告	按时提交（10%） 各模型在实验平台上的 测试代码（30%） 模型的优缺点分析（30%） 实验结果的可视化（20%） 答辩（10%）	40

从表 1 可以看出，在实验考核中将作业的按时提交列入了考核指标，旨在培养学生的契约精神，使其具备良好的执行力。

针对理论课的教学内容，实验课分为四次，包括决策树模型、回归模型、支持向量机模型以及神经网络模型。第一次实验课介绍通用实验框架的使用，包括实验环境配置、工具包配置、数据集、评价指标以及参数调优方法等，其余三次实验课只需要介绍实验框架中所用模型对应的 `scikit-learn` 中的类和方法。为了对比每个模型优缺点，在四次实验中均使用相同的数据集，回归任务使用房价预测数据集（`fetch_california_housing`），分类任务使用红酒分类数据集（`load_wine`）。具体授课内容安排如表 2 所示。

通用实验框架的构建是基于 `scikit-learn` 软件包进行构建的，其中包括数据集模块、机器学习模型模块、数据预处理模块、构建模型模块、评价指标以及模型预测结果模块。通用实验平台框架如图 1 所示。

表 2 授课内容安排

序号	实验中用到的模型	学时	实验内容
1	决策树和随机森林模型	2	实验环境配置 工具包介绍 (Numpy、Matplotlib、Pandas、scikit-learn) 数据集介绍 (load_boston、load_wine) 评价指标介绍 ( $R^2$ 、MSE、MAE; accuracy_score、Recall、F-score) scikit-learn中的决策树模型 (DecisionTreeRegressor、DecisionTreeClassifier) 和随机森林模型 (RandomForestRegressor、RandomForestClassifier) 参数调优方法 (GridSearchCV) 概述基础模型实现的流程
2	回归模型	2	scikit-learn中的回归模型 (LinearRegression、Ridge、Lasso、ElasticNet、LogisticRegression) 参数调优方法 (GridSearchCV) 概述基础模型实现的流程
3	支持向量机模型	2	scikit-learn中的支持向量机模型 (SVR、SVC) 参数调优方法 (GridSearchCV) 概述基础模型实现的流程
4	神经网络模型	2	增加手写数字识别数据集 (load_digit) scikit-learn中的神经网络模型 (MLPRegressor、MLPClassifier) 参数调优方法 (GridSearchCV) 概述基础模型实现的流程

表 3 回归模型的实验任务设计

<b>任务1: 预测任务</b>
数据集: 房价预测数据集 (load_boston)
实验内容: 分别使用线性回归 (LinearRegression)、岭回归 (Ridge)、套索回归 (Lasso) 以及弹性网络回归 (ElasticNet) 来预测房价
实验要求:
(1) 使用房价预测数据集 (fetch_california_housing) 的全部特征来预测房价, 分别使用 $R^2$ 、MSE、MAE指标来评估实验结果, 并写出预测模型的参数
(2) 对套索回归的参数调优 (使用GridSearchCV)
<b>任务2: 逻辑回归</b>
数据集: 红酒数据集 (load_wine)
实验内容: 使用逻辑回归 (LogisticRegression) 对红酒分类
实验要求:
(1) 使用红酒数据集 (load_wine) 的全部特征来预测红酒分类, 分别使用Accuracy、Recall以及F-Score指标来评估实验结果, 并写出分类模型的参数。
(2) 使用PCA方法预处理数据集并预测红酒分类, 与1) 中的结果进行比较。

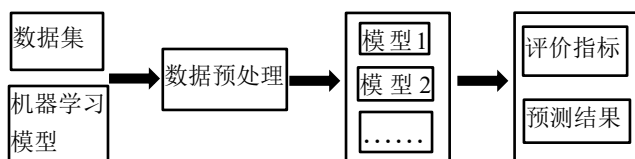


图 1 通用实验平台框架

在该通用实验平台框架给出了在机器学习实验中可选的数据集和机器学习算法, 为了便于学生全面掌

握模型之间的优缺点, 在数据集模块中给出多个数据集, 在本实验中给学生布置 2 个必选的数据集即房价预测数据集 (load\_boston) 和红酒数据集 (load\_wine), 这两个数据集的规模较小, 能满足在 2 学时的课堂时间完成实验任务, 其中, 房价预测数据集用于回归模型, 红酒数据集用于分类模型; 机器学习算法库中除了给出了实验课要求的决策树和随机森林模型、回归模型、支持向量机模型、神经网络模型外, 还包括了聚类模型、贝叶斯模型等。

数据预处理模块是一个可选模块，在数据预处理模块中给出了数据转换方法、主成分分析方法等数据预处理方法、构建模型模块用于存储使用不同数据集和机器学习方法所构建的模型；评价指标模块给出回归和分类任务的不同评价指标；模型预测结果模块用于存储和显示模型的预测结果，便于学生对比模型的预测结果并进行分析和可视化。学生通过使用通用实验框架有利于后续综合实验报告的编写，也有利于查看在相同数据集上不同模型的评价结果。

## 4 机器学习实验课教学设计实例

以回归模型的实验为例介绍教学设计内容，回归模型实验是第2次实验课，学时为2学时。在第1次

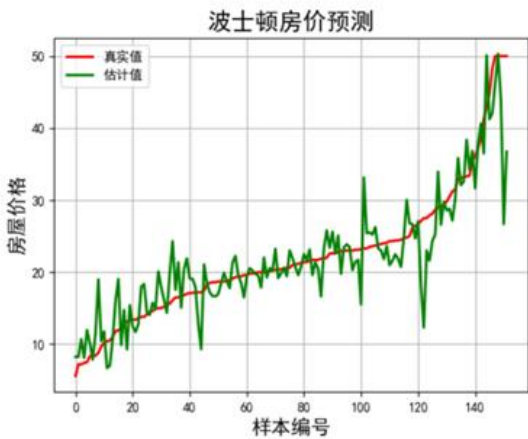
实验课中已经给出这4次实验中所用到的开发环境、工具包以及数据集，本次回归模型的实验设计主要介绍回归模型在通用实验平台中所用到的方法对应的scikit-learn中所用到的类、参数调优方法。回归模型的实验任务设计如表3所示。

### 4.1 回归模型所用到的类

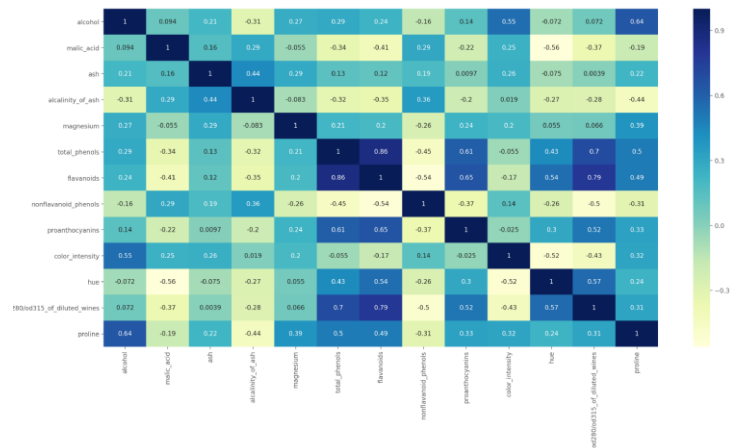
基于回归模型理论课的授课内容，将回归模型实验设置为回归任务和分类任务，回归任务分别使用线性回归、岭回归、套索回归以及弹性网络回归模型实现；分类任务使用逻辑回归模型实现。每个模型对应的scikit-learn的类定义如表4所示。

表4 模型对应的类定义

模型名	类定义
线性回归	<code>class sklearn.linear_model.LinearRegression(*, fit_intercept=True, copy_X=True, n_jobs=None, positive=False)</code>
岭回归	<code>class sklearn.linear_model.Ridge(alpha=1.0, *, fit_intercept=True, copy_X=True, max_iter=None, tol=0.0001, solver='auto', positive=False, random_state=None)</code>
套索回归	<code>class sklearn.linear_model.Lasso(alpha=1.0, *, fit_intercept=True, precompute=False, copy_X=True, max_iter=1000, tol=0.0001, warm_start=False, positive=False, random_state=None, selection='cyclic')</code>
弹性网络回归	<code>class sklearn.linear_model.ElasticNet(alpha=1.0, *, l1_ratio=0.5, fit_intercept=True, precompute=False, max_iter=1000, copy_X=True, tol=0.0001, warm_start=False, positive=False, random_state=None, selection='cyclic')</code>
逻辑回归	<code>class sklearn.linear_model.LogisticRegression(penalty='l2', *, dual=False, tol=0.0001, C=1.0, fit_intercept=True, intercept_scaling=1, class_weight=None, random_state=None, solver='lbfgs', max_iter=100, multi_class='auto', verbose=0, warm_start=False, n_jobs=None, l1_ratio=None)</code>



(1)



(2)

图2 作业中的部分可视化分析结果

## 4.2 参数调优方法

在回归模型的实验中，涉及到的模型较多，每个模型的可调参数也不多，最基本的线性模型是涉及调节参数最少的模型。在该部分采用启发式的方式引导学生调整参数来构建模型。例如，在套索回归中，调整  $\alpha$  和  $\max\_iter$  参数观察对预测结果的影响，代码如下所示。

```
#设置alpha, max_iter的可能取值
params={'alpha':[0.001,0.01,0.1,1.0,10],'max_iter':[100,
1000,3000,5000,10000]}
#设置参数
lasso=Lasso().fit(x_train,y_train)
gsv=GridSearchCV(lasso,params,cv=6).fit(x_train,y_train)
```

## 5 学生成果物展示

实验课结束后共收集到86份作业，综合评分后，本次回归模型作业的最高分10分的作业有19份，最低分为6分，共7份。作业中的部分可视化分析结果如图2所示。

由于篇幅有限这里只显示了学生作业中的2张可视化的效果图，图2中第1张图片显示的弹性网络回归模型预测房价的结果，第2张图片是用混淆矩阵的方式显示每个数据集中每个特征之间的关系。

## 5 结束语

针对机器学习课程实验课教学存在的教材选择缺乏针对性、实验内容之间缺乏衔接以及课程形式缺乏多样性三方面的不足，本文结合机器学习理论课所学的内容设计了4次实验课，并在文中列举了1次实验课的实施过程和效果。86名大三同学通过此次机器学习课程实验课的学习和实践，在灵活运用理论课相关知识的基础上培养了其自主学习能力，并激发了学生对模型优化的探索精神。

### 参考文献

- [1] 叶景贞,康盛,方梓荷.机器学习课程的教学实践[J].电子技术(上海),2021(9):2.
- [2] 李阳.机器学习课程的设计与教学模式实践[J].电子技术(上海),2022.
- [3] 王敏,罗婧雯,刘军,等.基于朴素贝叶斯的机器学习实验教学设计与实践[J].中国新通信,2021,23(6):2.
- [4] 徐国艳,周彬,刘磊,等.基于百度AI Studio的机器学习课程教学设计与实践[J].计算机教育,2021,000(006):100-104,109.
- [5] 黄美燕.基于医工结合的“机器学习”课程教学设计研究[J].教师,2023,(26).