

**《专业综合实践II》课程考核**

|  |  |
| --- | --- |
| **题 目：** | 基于机器学习方法的手写数字识别研究 |
|  |  |
| **学 院：** | 数据科学学院 |
| **专 业：** | 数据科学与大数据技术 |
| **年级班别：** | 2021级（1）班 |
| **学 号：** | 421470133 |
| **学生姓名**： | 刘小惠 |

**提交日期：** 2024 **年** 11 月

**目 录**

[1. 绪论 1](#_Toc15308)

[1.1. 研究背景 1](#_Toc8935)

[1.2. 研究目的 1](#_Toc4623)

[1.3. 研究方法 1](#_Toc18141)

[2. 基于机器学习方法的图像分类算法 2](#_Toc4749)

[2.1. 机器学习方法 2](#_Toc24160)

[2.2. 常用算法 3](#_Toc26286)

[2.2.1. 卷积神经网络（CNN） 3](#_Toc24865)

[2.2.2. 支持向量机（SVM） 3](#_Toc29598)

[2.2.3. K最近邻（KNN） 3](#_Toc17113)

[2.2.4. 循环神经网络（RNN） 3](#_Toc23687)

[2.2.5. 决策树（Decision Tree） 3](#_Toc26381)

[3. 图像收集和预处理 3](#_Toc14264)

[3.1. 数据集介绍 3](#_Toc20528)

[3.2. 图像预处理 4](#_Toc5583)

[3.3. 分类器 4](#_Toc13150)

[3.3.1. 决策树分类器 4](#_Toc14323)

[3.3.2. 支持向量机分类器 4](#_Toc12459)

[3.3.3. 近邻分类器 5](#_Toc24269)

[3.4. 参数 5](#_Toc18552)

[3.5. 数据可视化 5](#_Toc14453)

[3.6. 训练与优化策略 7](#_Toc1523)

[4. 结果与讨论 7](#_Toc15166)

[4.1. 实验结果 7](#_Toc11809)

[4.2. 结果分析 7](#_Toc17717)

[5. 结论与展望 8](#_Toc23150)

[5.1. 主要结论 8](#_Toc20055)

[5.2. 未来工作展望 8](#_Toc20239)

[参考文献 8](#_Toc4605)

[得 分 9](#_Toc10918)

# 绪论

## 研究背景

手写数字识别是计算机视觉和模式识别领域中的一个经典问题，它涉及到从不同个体书写的数字图像中准确识别出相应的数字。这一技术在邮政编码分拣、银行支票处理、表单数据录入等多个实际应用场景中发挥着重要作用。传统的手写数字识别方法大多依赖于特征工程，即人工设计和提取图像的特征。这些方法通常依赖专业人员的经验和知识来选择和设计对分类有帮助的特征。然而，这类手工方法不仅耗时耗力，而且一旦遇到复杂的实际问题，特征的选择和设计将变得极为困难。因此，如何利用机器学习技术自动并高效地完成手写数字的识别，成为了研究的重要课题。

## 研究目的

本文的目的是探索和比较不同的机器学习方法在手写数字识别任务上的性能表现，特别是K近邻（KNN）、支持向量机（SVM）和决策树这三种算法。通过对比这些算法在该任务上的优势和局限，可以为实际应用中算法的选择提供依据，并推动手写数字识别技术的发展。

## 研究方法

首先，我们使用公开的MNIST数据集作为实验的基础。这个数据集包含了6000个训练样本和1000个测试样本，每个样本都是28x28像素的灰度图像，代表一个0到9之间的手写数字。

对于KNN算法，我们直接计算测试样本与训练集中每个样本的欧式距离，然后选取距离最近的K个邻居，根据这K个邻居的标签进行投票来决定测试样本的类别。K值的选择将通过交叉验证来确定。

在SVM的应用中，我们将采用非线性的SVM，使用径向基函数（RBF）作为核函数。通过调整惩罚参数C和核函数的参数gamma来优化SVM的性能。参数的优化同样采用交叉验证的方法。

对于决策树，我们使用ID3或CART算法构建决策树。为了提高模型的泛化能力，我们引入剪枝策略来避免过拟合。

# 基于机器学习方法的图像分类算法

## 机器学习方法

机器学习是一种人工智能领域的技术，它赋予计算机系统通过经验来改进的能力。与传统的编程方法不同，机器学习依赖于数据和算法，使计算机能够从数据中学习并做出决策或预测，而无需进行明确的编程指令。

在机器学习方法中，通常分为三个主要类别：监督学习、无监督学习和强化学习。监督学习涉及使用标记好的数据集来训练模型，目标是让模型学会将输入映射到正确的输出。例如，分类问题（如邮件是垃圾邮件还是非垃圾邮件）和回归问题（如预测房价）都是监督学习的应用。在训练过程中，模型尝试最小化其预测与实际标签之间的差异。常用的监督学习算法包括决策树、随机森林、支持向量机、神经网络等。

无监督学习则处理未标记的数据。它的目标是发现数据中的结构，比如聚类（将相似项分组在一起）或降维（减少数据的复杂性，同时尽量保留重要信息）。典型的无监督学习算法有 k-均值聚类、主成分分析（PCA）和自组织映射等。

强化学习不同于前两者，它不依赖预先存在的数据集来学习，而是通过智能体与环境的交互来进行学习。智能体根据其所采取的行动获得奖励或惩罚，目标是找到一种策略，最大化长期累积奖励。强化学习被广泛应用于游戏、机器人控制等领域。著名的算法包括 Q-learning 和深度强化学习算法，如深度 Q 网络（DQN）。

机器学习方法还包括深度学习，这是一种特殊的机器学习技术，它使用多层神经网络来学习复杂的模式和特征。深度学习在图像识别、语音识别和自然语言处理等领域取得了显著的成功。

机器学习方法的应用范围广泛，从自动驾驶汽车、医疗诊断到股票市场分析、推荐系统等等。随着技术的发展，机器学习方法不断进步，包括自动化机器学习（AutoML）的出现，它旨在自动化设计机器学习模型的过程，降低机器学习门槛，使更多非专业人员也能利用这一强大工具。

总之，机器学习方法是现代科学的重要组成部分，它为解决复杂问题提供了强大的工具，同时也带来了众多挑战，如数据隐私、模型解释性和伦理问题。随着技术的不断演进，机器学习将持续影响我们的生活和工作方式。

## 常用算法

### 卷积神经网络（CNN）

卷积神经网络（CNN）是目前最优的网络结构之一，特别适合于图像分类任务。在手写数字识别中，CNN通过卷积层、池化层和全连接层逐步提取图像特征，最终实现高精度的识别。例如，使用PyTorch框架搭建的CNN模型可以达到99%的准确率。

### 支持向量机（SVM）

支持向量机（SVM）是一种监督学习的方法，可以用于分类和回归分析。它通过找到最佳的超平面来划分不同的类别。在手写数字识别中，SVM也被广泛应用，但通常不如CNN表现出色。

### K最近邻（KNN）

K最近邻（KNN）算法通过计算测试样本与训练样本之间的距离，选择最近的K个样本进行投票，以确定测试样本的类别。在手写数字识别中，KNN算法简单易实现，但其性能通常不如CNN和SVM。

### 循环神经网络（RNN）

循环神经网络（RNN）适用于处理序列数据，如时间序列或自然语言处理。在手写数字识别中，RNN可以通过逐步处理每个数字的笔画来提高识别效果。

### 决策树（Decision Tree）

**决策树**（Decision Tree）**是一种基于树形结构的决策模型。 它通过一系列的决策节点和分支，对数据进行分类或回归**。 每个决策节点代表一个特征属性上的判断条件，分支则代表不同的判断结果，最终每个叶节点代表一个类别或数值预测。

# 图像收集和预处理

## 数据集介绍

本研究选用了广泛应用于手写数字识别领域的MNIST数据集。该数据集包含了60000个训练样本和10000个测试样本，每个样本都是28x28像素的灰度图像，代表了0到9的手写数字。MNIST数据集的多样性和规模使其成为评估手写数字识别算法性能的理想选择。

## 图像预处理

### 在机器学习中，图像预处理步骤通常包括图像尺寸调整、归一化处理、数据增强等。而归一化是其中重要的一环，其目的是将图像的像素值缩放到一个特定的范围，通常是0到1，有助于模型训练的收敛速度和提高模型的表现稳定性。

图像预处理的步骤：

（1）图像尺寸调整：为了确保输入网络的图像有统一的尺寸，需要调整图像大小以匹配模型的输入要求。例如，某些模型可能需要 224×224 或 227×227 的输入图像尺寸。

（2）归一化处理：归一化是将图像素值规范化到一个给定的范围，常见的是[0, 1]。这样做可以减少像素值范围对于模型训练时权重更新的影响。

（3）数据增强：通过旋转、翻转、裁剪、色彩变换等手段增加数据的多样性，这有助于提升模型对于新数据的泛化能力。

（4）去噪和清晰度改善：使用各种滤波器和技术减少图像中的噪声并提高图像质量，有助于提升模型对图像特征的捕捉能力。

（5）直方图均衡化：用于改善图像的对比度，使图像中的特征更加明显。

## 分类器

在本研究中，本人使用看以下三类分类器。

### 决策树分类器

决策树（Decision Tree）是一种监督学习算法，非常适合解决分类问题。它通过构建树状图来对数据进行分类，可以精确地对类别进行排序。决策树分类器的优点是简单直观，易于理解和实现。

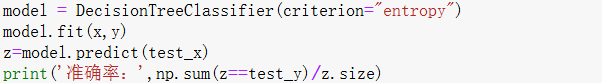


图1.1 决策树（Decision Tree）的代码

### 支持向量机分类器

支持向量机 (SVM) 是一种常用的分类算法，它通过找到能够将不同类别的数据点最大化分隔的决策边界来实现分类。SVM 的优点是能够处理高维数据，并且在处理小样本、非线性数据和异常值方面表现良好。

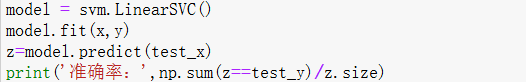


图1.2 支持向量机 (SVM)的代码

### 近邻分类器

近邻分类器是一种基于实例的学习算法，它通过将新数据点与已知数据点进行比较来对数据进行分类。最常见的近邻算法是 k-最近邻 (KNN)，它通过找到与新数据点最接近的 k 个数据点来决定分类。近邻分类器的优点是简单、直观，但是计算量大，适用于小样本数据集。

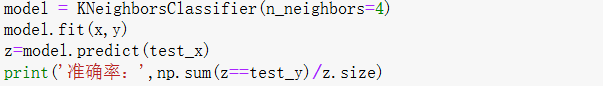


图1.3 K最近邻（KNN）的代码

## 参数

手写数字识别技术中使用的参数主要包括：

（1）预处理参数：如二值化阈值、归一化尺寸等。

（2）特征提取参数：如边缘检测算子、纹理描述子等。

（3）特征选择参数：如特征选择算法、特征维度等。

（3）分类器参数：如支持向量机的核函数、神经网络的层数和神经元数量等。

## 数据可视化

在机器学习领域中，数据可视化具有非常重要的作用。首先，数据可视化可以帮助我们更好地理解数据，发现数据中的规律和趋势。这对于后续的算法设计和模型训练非常重要。其次，数据可视化可以让我们更好地了解机器学习算法的性能和表现，从而更好地优化算法和提高模型准确率。此外，数据可视化还可以帮助我们更好地进行交互式分析和探索性分析，发现新的信息和知识，为决策提供有力的支持。

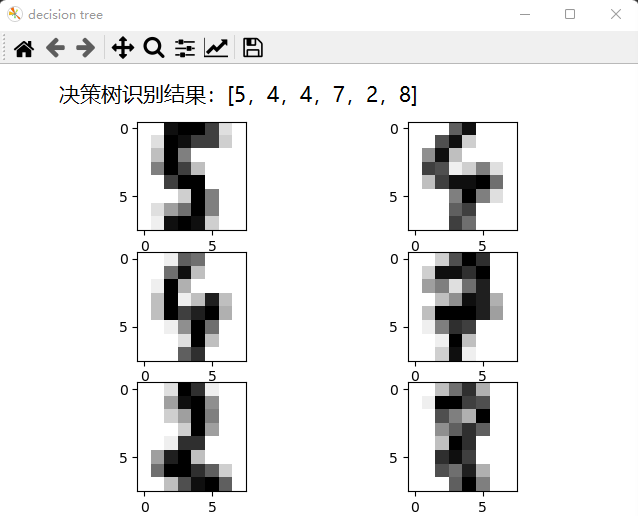


图2.1 **决策树**（Decision Tree）的识别结果

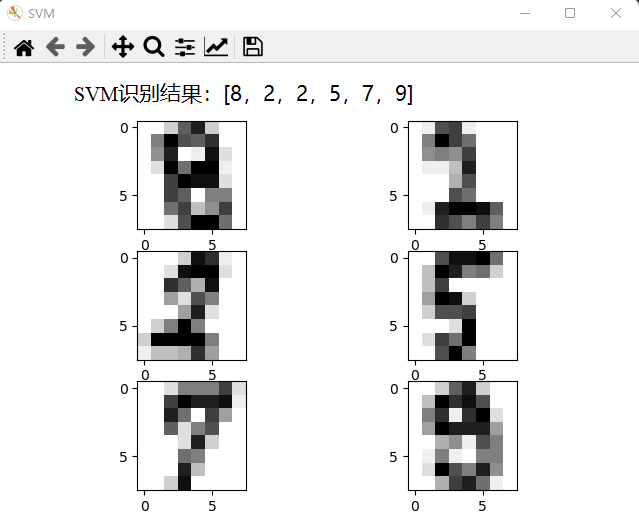


图2.2 支持向量机（SVM）的识别结果

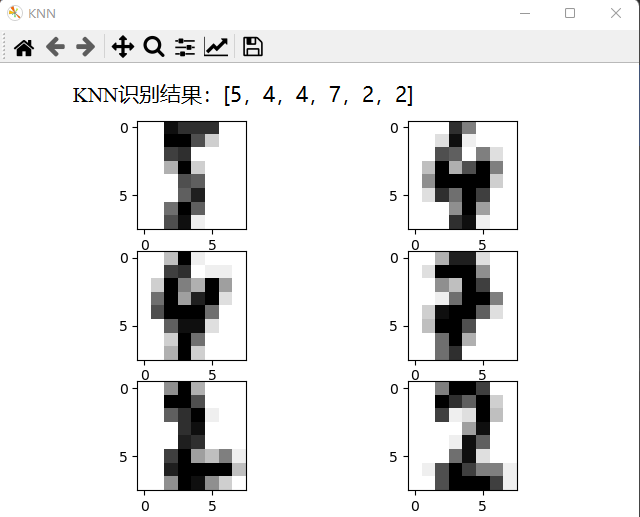


图2.3 K最近邻（KNN）的识别结果

可视化通常是对模型性能的一个补充，而不是唯一的评估手段，因此在实际应用中还需要结合其他评估指标（如准确度、召回率、f1分数等）来全面评价模型的表现。

## 训练与优化策略

## 在模型训练阶段，采用了多种优化策略以提高学习效率和模型性能。使用了Adam优化器进行参数更新，它能够自适应地调整学习率。为了防止过拟合，引入了dropout技术随机丢弃一部分神经元。此外，还应用了批量归一化技术来加速训练过程并提高模型的稳定性。通过这些策略的应用，确保了模型能够在保持较高准确率的同时，也具有较好的泛化能力。

# 结果与讨论

## 实验结果



图3.1 **决策树**（Decision Tree）的准确率



图3.2支持向量机（SVM）的准确率



图3.3 K最近邻（KNN）的准确率

在本研究中，通过结果可以看出KNN分类器表现的效果比较理想，准确率高达97.3%；虽然SVM的训练阶段成本较高，但效果一般；决策树分类器出现了过拟合，需要进行预剪枝来解决过拟合，可以通过设置最大树深以及其他超参。

## 结果分析

通过对实验结果的分析，可以看出采用的三个模型能够更有效地从手写数字图像中提取和学习到有用的特征。总的来说，SVM在非线性可分数据的处理上具有优势，而决策树在模型解释性和可视化方面表现更佳。KNN虽然在处理大规模数据时表现出一定的局限性，但在特定场景下仍能提供可靠的识别能力。在实际应用中，选择合适的算法需要考虑数据的性质、问题的复杂性以及计算资源等多个因素。这些结果不仅展示了机器学习方法在手写数字识别任务中的应用潜力，也为未来相关研究提供了有价值的参考。

# 结论与展望

## 主要结论

本研究通过结合**决策树**（Decision Tree）、支持向量机（SVM）和K最近邻（KNN）的方法，显著提高了手写数字识别的准确率和效率。实验结果表明，所使用的三个模型在MNIST数据集上最高取得了97.3%的高识别准确率，这一成果证实了深度学习在手写数字识别领域的有效性，并为进一步的研究和应用提供了坚实的基础。基于机器学习方法的手写数字识别已经取得了显著的进展，尤其是深度学习技术在此领域表现出色。通过使用决策树、SVM和KNN等算法，可以有效地提高手写数字识别的准确率和效率。实际应用中，选择合适的框架和模型架构是关键，这将直接影响到系统的性能和可靠性。

## 未来工作展望

未来的工作将集中在以下几个方面：首先，探索更多的网络结构和深度学习技术，以进一步提升识别性能；其次，将研究扩展到更复杂的数据集和实际应用场景，如不规则手写文本的识别；最后，考虑到计算资源的限制，研究更高效的模型压缩和加速技术，以便在资源受限的设备上部署。通过这些努力，期望将手写识别技术推向更高的水平，满足日益增长的实际应用需求。

# 参考文献

[1]周志华.机器学习[ M ] .清华大学出版社，2016.

[2]赵鹏，王晓东，&张学工.支持向量机理论与方法综述J1 .自动化学报，2002.28 ( 6 )：788-794.

[3]张学工，赵鹏.支持向量机及其在模式识别中的应用J1 .模式识别与人工智能.