

**手写数字图像分类识别**

|  |  |
| --- | --- |
| **题 目：** | 手写数字图像分类识别 |
| **学 院：** | 人工智能学院 |
| **专 业：** | 数据科学与大数据技术 |
| **年级班别：** | 2021级数据科学与大数据技术1班 |
| **学 号：** | 421470101 |
| **学生姓名**： | 艾伟民 |

**提交日期：** 2024 **年** 11 月

# 摘要

本文提出一种基于深度学习的手写0到5数字识别算法，旨在实现快速准确的识别和分类0到5的手写数字。手写体数字识别是文字识别中的个研究课题。 由于识别类型较少，在实际生活中有深远的应用需求，一直得到广泛的重视。近年来随着计算机技术和数字图像处理技术的飞速发展，数字识别在电子商务，机器自动输入等场合已经获得成功的实际应用。

对于手写体数字识别，目前已经发展了很多种方法，它们是基于神经网络算法，基于笔划特征的算法，基于遗传算法，基于小波变换算法，基于傅立叶变换算法，基于支持向量机算法和基于模板匹配算法等等。由于手写体数字识变体极多，对各类字体的数字识别特别是脱机手写数字识别仍然处于在发展阶段，识别效果仍然不够理想。因此，研究简单高效的手写数字识别依然是一个重要的研究方向。本文通过图像预处理和数字特征提取，基于神经网络的判别方法，并结合使用Matlab工具箱中提供的人工神经网络函数设计了一种手写数字识别的新方法。实验表明，该方法可以获得较好的识别率。

该研究拍摄了0到5的手写数字图像，每个数字各有120的图像数据，一共有6个图像类别，并在实验开始前就已经分类好，通过数据清洗后转移到一个新的文件夹里进行实验，使用SVM，随机森林，多层感知器这三种分类器进行实验，对比这三种分类器的精确度，f1分数和召回率找到表现最好的分类器。

**关键词：**数字图像识别，SVM，随机森林，多层感知器。

# 目录

[摘要 1](#_Toc6932)

[1绪论 3](#_Toc30253)

[1.1研究背景与研究意义 3](#_Toc30950)

[1.2 SVM、随机森林、多层感知器的介绍 3](#_Toc21598)

[2 手写数字图像的数据描述 4](#_Toc30716)

[2.1 数据来源 4](#_Toc2113)

[2.2 数据类型 4](#_Toc20100)

[3 数据处理 5](#_Toc30885)

[3.1 数据清理 5](#_Toc19457)

[3.2数据归类 6](#_Toc11351)

[4 机器学习 7](#_Toc6593)

[4.1 SVM分类器 7](#_Toc3373)

[4.2 随机森林分类器 8](#_Toc2847)

[4.3 多层感知器分类器 8](#_Toc10009)

[4.4 机器学习代码 9](#_Toc30311)

[5数据可视化 1](#_Toc12473)3

[6机器学习结论 1](#_Toc4926)4

[6.1 机器学习的结论 1](#_Toc22706)4

# 1绪论

## 1.1研究背景与研究意义

数字图像识别技术在现代社会中扮演着重要的角色，它被广泛应用于安全监控、图像搜索、自动驾驶汽车等领域。传统的图像识别方法主要依赖于特征工程，然而，在面对复杂的图片数据时，传统方法常常面临挑战。随着机器学习尤其是深度学习的发展，图片识别的准确性得到了显著提升。

本研究旨在比较SVM，随机森林，多层感知器等分类方法在数字图像识别中的应用。通过实验，我们旨在探究这些方法在数字图像识别任务中的性能差异，以找到最适合该任务的分类方法。

本文的主要方法是：首先，收集一个包含多种数字样式的图像数据库。然后，分别使用SVM，随机森林，多层感知器设计三个图像分类器。通过训练和优化这三个分类器，我们期望在测试集上获得较高的识别准确率。最后，我们将评估所提出方法在识别数字图像中的性能。

本研究的意义在于为数字图像识别提供一种新的方法，并为后续研究提供参考。对于实际应用，如自动读取银行支票上的数字、对安全监控图片中的数字进行识别等，本研究的成果具有很高的实用价值。

## 1.2 SVM、随机森林、多层感知器的介绍

支持向量机（Support Vector Machine, SVM）是一类按监督学习方式对数据进行二元分类的广义线性分类器，其决策边界是对学习样本求解的最大边距超平面。SVM使用铰链损失函数计算经验风险并在求解系统中加入了正则化项以优化结构风险，是一个具有稀疏性和稳健性的分类器 。SVM可以通过核方法进行非线性分类，是常见的核学习方法之一。

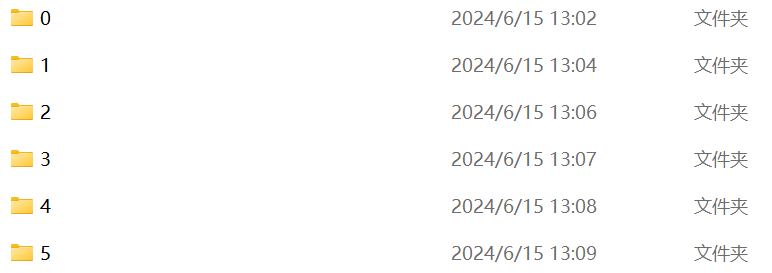
随机森林是一个包含多个决策树的分类器，并且其输出的类别是由个别树输出的类别的众数而定。是一种包含很多决策树的分类器，既可以用于处理分类和回归问题，也适用于降维问题。其对异常值与噪音也有很好的容忍，相较于决策树有着更好的预测和分类性能。

多层感知器（MLP，Multilayer Perceptron）是一种前馈人工神经网络模型，其将输入的多个数据集映射到单一的输出的数据集上。

# 2 手写数字图像的数据描述

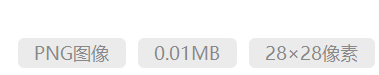
## 2.1 数据来源

本实验用手机拍摄0到5的手写数字图像，并将拍摄的图像按数字类别进行分类。



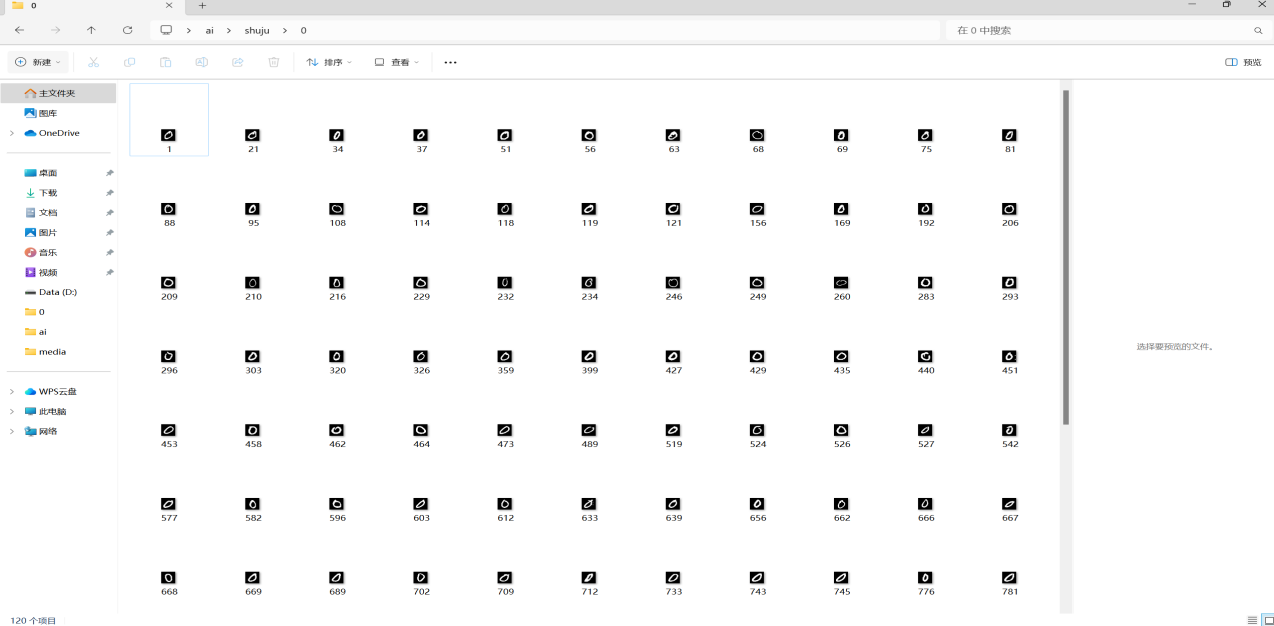
## 2.2 数据类型

本实验的数据是手写数字的拍摄图像，格式是PNG图像，28\*28的像素



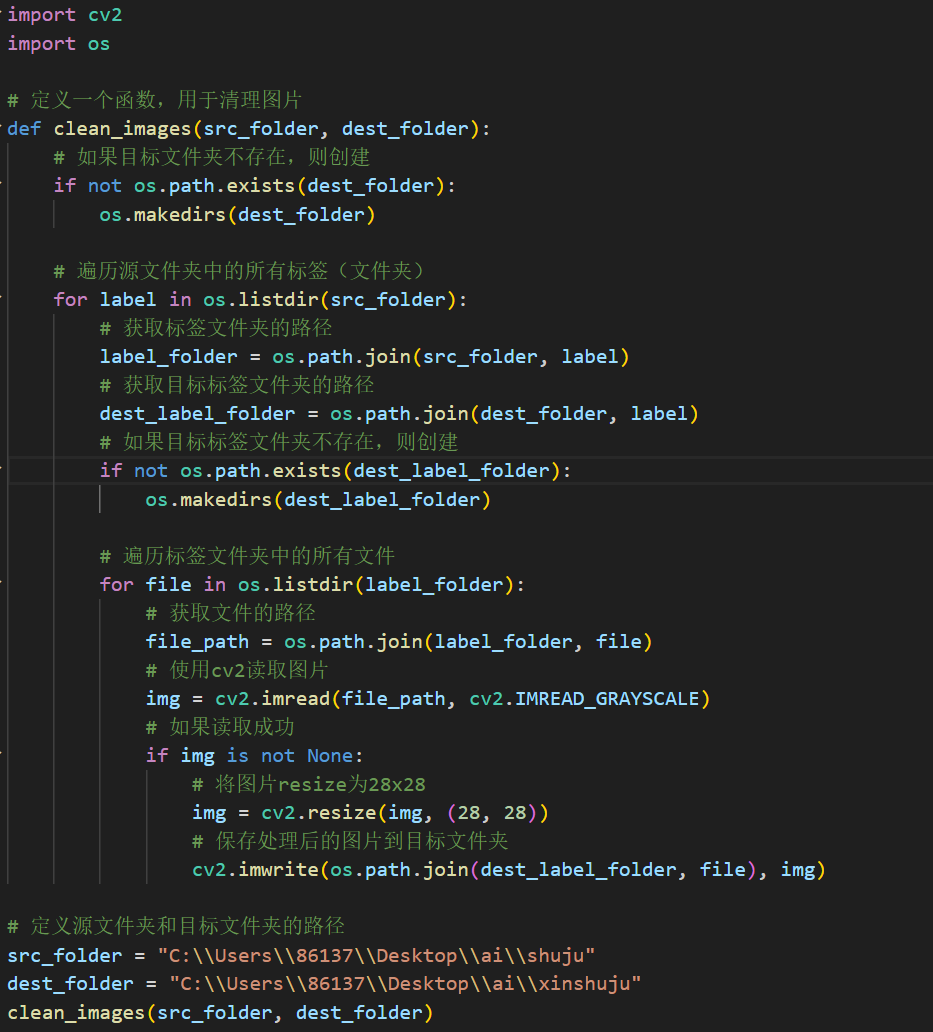
2.3数据规模

本实验的数据一共有720个，一共有6个不同的数据类别，分别为0，1，2，3，4，5的手写数字图像，每组图像都有120个数据



# 3 数据处理

## 3.1 数据清理



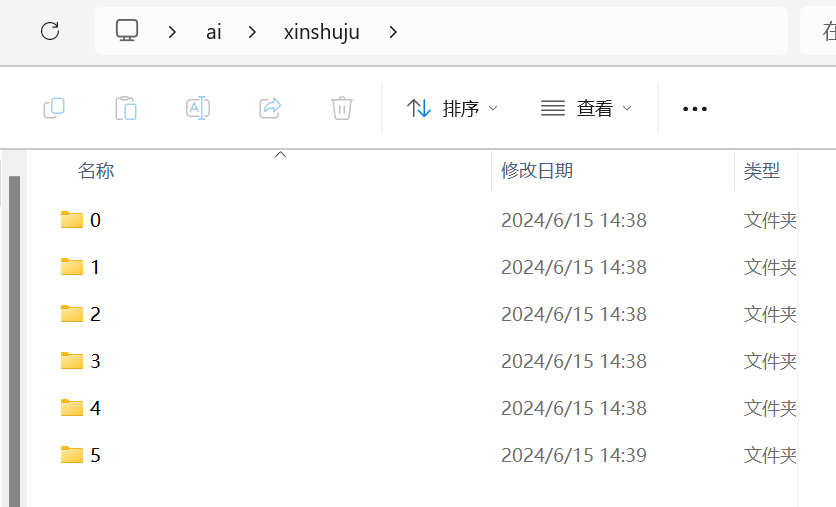
这段代码的目的是清理图片，将图片转换为灰度图并调整大小为28x28。导入所需的库：`cv2`用于处理图片，`os`用于操作文件和目录。定义一个名为`clean\_images`的函数，该函数接受两个参数：`shuju`（源文件夹路径）和`xinshuju`（目标文件夹路径）。如果目标文件夹不存在，则创建目标文件夹。遍历源文件夹中的所有标签（文件夹）。对于每个标签文件夹，获取标签文件夹的路径，并创建目标标签文件夹（如果尚不存在）。遍历标签文件夹中的所有文件。对于每个文件，获取文件的路径，并使用`cv2.imread`以灰度模式读取图片。如果图片读取成功（即图片不是`None`），则将图片调整为大小28x28，并将处理后的图片保存到目标文件夹。最后，定义源文件夹和目标文件夹的路径，并调用`clean\_images`函数以执行清理操作。

## 3.2数据归类

本实验的原始数据集（0到5的手写数字图像）在进行收集的时候已经进行了归类，并且进行数据清理后把清理后的数据放在了新的文件夹里，新的文件夹也是已经归类好了。

如图：





原始数据位于“shuju”的文件夹里，清洗后的数据位于“xinshuju”里。

# 4 机器学习

## 4.1 SVM分类器

支持向量机（Support Vector Machine，SVM）是一种用于二类分类问题的学习算法。SVM的基本思想是找到一个超平面，该超平面可以将数据集中的点尽可能地分开。这个超平面就是所谓的“支持向量”，而SVM的目标就是找到一个最佳的超平面，使得支持向量之间的间隔最大。

SVM是一种基于统计学习理论的分类方法，它通过在特征空间中找到一个超平面，使得正负样本的间隔最大。SVM适用于线性分类问题，对于非线性分类问题，可以通过使用核函数将数据映射到高维空间中，从而实现线性分类。

SVM的主要优点包括：

1. 非线性分类能力：SVM可以通过使用核函数将数据映射到高维空间中，从而实现非线性分类。

2. 训练速度快：SVM的训练速度通常比其他分类方法更快。

3. 模型简单：SVM的实现简单，易于理解和实现。

4. 结果准确率较高：SVM通常能够获得较高的分类准确率。

SVM的主要缺点包括：

1. 对参数敏感：SVM的参数对结果影响较大，不同的参数设置可能会导致不同的分类结果。

2. 无法处理多分类问题：SVM仅适用于二类分类问题，对于多类分类问题，需要使用多分类SVM，例如多类逻辑回归（Multi-class Logistic Regression）或多类感知器（Multi-class Perceptron）。

3. 特征选择困难：SVM需要对特征进行选择，以提高分类性能。特征选择困难可能会导致SVM的性能下降。

4. 训练数据不足：SVM对于训练数据的需求较高，如果训练数据不足，可能会导致过拟合问题。

总之，SVM是一种基于统计学习理论的分类方法，具有非线性分类能力、训练速度快、模型简单和结果准确率较高的优点，但在参数敏感、无法处理多分类问题、特征选择困难等方面存在一些缺点。在实际应用中，需要根据具体问题进行权衡和优化。

## 4.2 随机森林分类器

随机森林（Random Forest）是一种集成学习方法，它采用多棵树来完成分类任务。这些树被称为决策树，因为它们类似于人类选择的决策树。随机森林通过构建多个决策树，然后对这些树的结果进行投票或平均来完成分类。

随机森林的主要优点包括：

1. 随机森林算法对数据中的噪声和异常值不敏感，具有很好的抗噪能力。

2. 随机森林算法具有很好的特性，即使在一些特征与目标变量不相关的情况下，算法仍然可以表现良好。

3. 随机森林算法可以产生较高的准确率。

随机森林的主要缺点包括：

1. 随机森林算法在训练过程中会产生大量的决策树，占用较大的存储空间。

2. 随机森林算法的可解释性较差，难以解释其分类结果。

## 4.3 多层感知器分类器

多层感知器（MLP）是一种前馈神经网络，它由多层神经元组成，可以用于分类和回归分析。MLP可以通过激活函数实现非线性映射，因此可以用于处理非线性数据。

MLP的主要优点包括：

1. MLP可以用于处理非线性数据，具有很好的通用性。

2. MLP可以通过调整网络结构（如增加隐藏层数、神经元数量等）来提高模型性能。

3. MLP在训练过程中可以自动提取特征，无需手动特征工程。

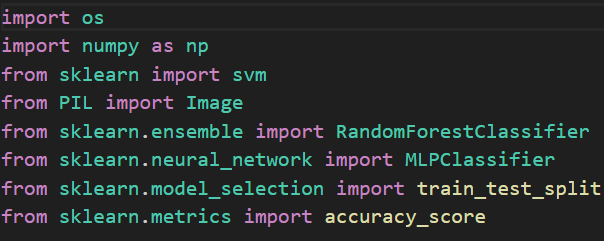
MLP的主要缺点包括：

1. MLP对超参数敏感，如学习率、隐藏层数、神经元数量等，较难调整。

2. MLP可能出现过拟合现象，特别是在数据量较小的情况下。

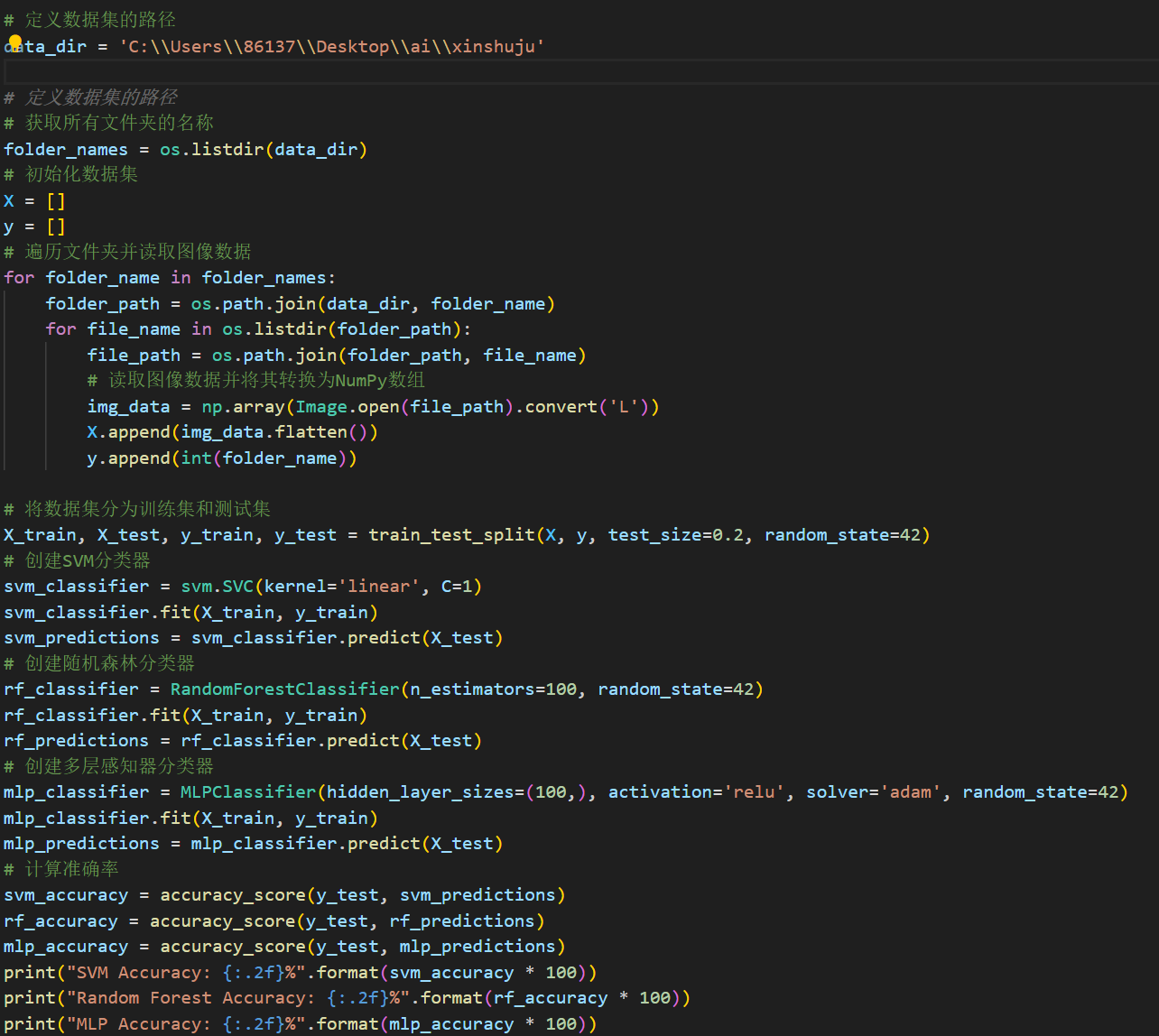
## 4.4 机器学习代码

4.4.1需要用到的库



这段代码首先导入了几个常用的Python库，包括os、numpy、scikit-learn库中的svm（支持向量机）、PIL（Python Imaging Library，用于处理图像）、RandomForestClassifier（随机森林分类器）、MLPClassifier（多层感知器分类器）等。接下来，代码中定义了几个变量，包括Image.open()用于打开图像文件，train\_test\_split()用于将数据集划分为训练集和测试集，accuracy\_score()用于计算模型预测的准确性。这段代码的主要用途是实现一个图像分类器，使用支持向量机、随机森林和多层感知器等分类器进行图像分类，并计算分类的准确性。

4.4.2 创建分类器，计算准确率



这段代码主要完成以下任务：

1. 定义数据集的路径，这里的数据集是一个包含多个文件夹的桌面文件夹，每个文件夹代表一个类别，文件夹内包含该类别的图像数据。

2. 获取所有文件夹的名称，并将其存储在`folder\_names`列表中。

3. 初始化数据集`X`和`y`，`X`用于存储图像数据，`y`用于存储图像对应的类别标签。

4. 遍历文件夹并读取图像数据，将图像数据转换为NumPy数组，并将其添加到`X`列表中，同时将类别标签添加到`y`列表中。

5. 将数据集分为训练集和测试集，其中测试集占20%。

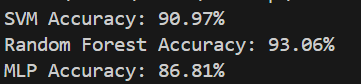
6. 创建SVM分类器，并使用训练集对其进行训练，然后使用测试集进行预测，并将预测结果存储在`svm\_predictions`中。

7. 创建随机森林分类器，并使用训练集对其进行训练，然后使用测试集进行预测，并将预测结果存储在`rf\_predictions`中。

8. 创建多层感知器分类器，并使用训练集对其进行训练，然后使用测试集进行预测，并将预测结果存储在`mlp\_predictions`中。

9. 计算SVM、随机森林和多层感知器的准确率，并将结果打印出来。

结果：



这段代码的运行结果给出了使用三种不同分类器（SVM、随机森林、多层感知器）对图像数据进行分类的准确率。

- SVM Accuracy: 90.97%

SVM（支持向量机）分类器的准确率是90.97%。这意味着在测试集中，SVM分类器正确分类的图像数量占测试集中图像总数的90.97%。

- Random Forest Accuracy: 93.06%

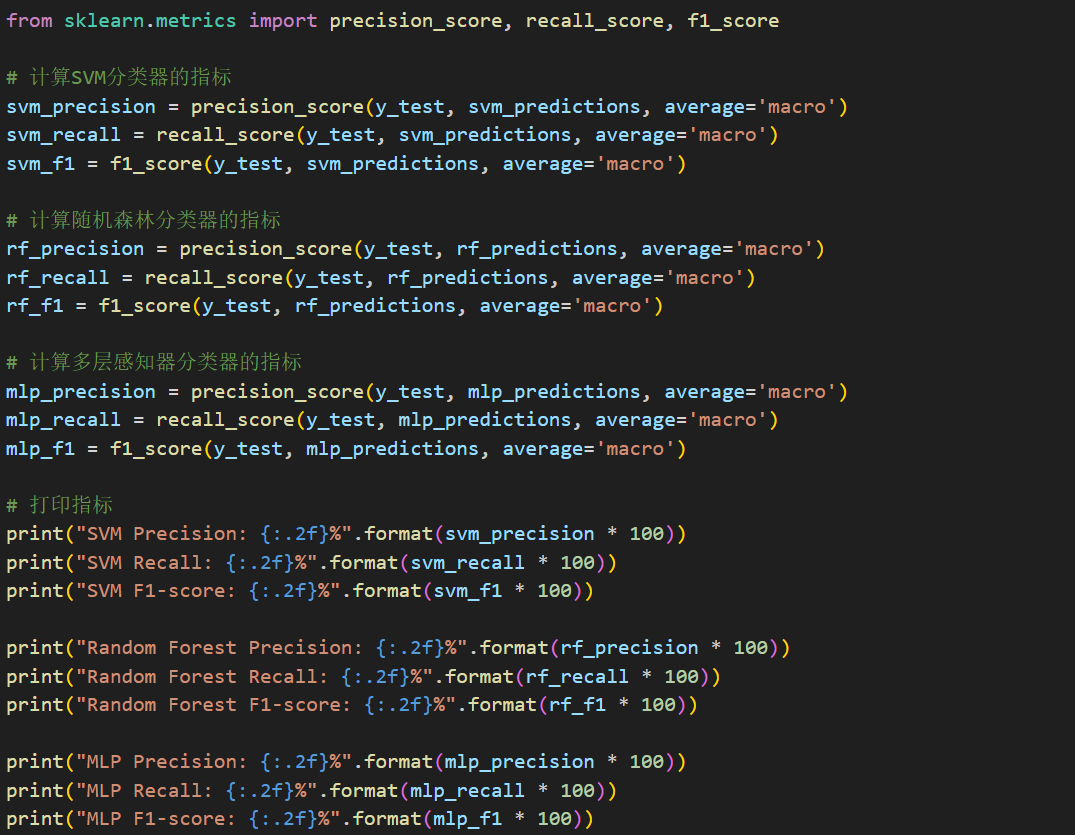
随机森林分类器的准确率是93.06%。这意味着在测试集中，随机森林分类器正确分类的图像数量占测试集中图像总数的93.06%。

- MLP Accuracy: 86.81%

多层感知器（MLP）分类器的准确率是86.81%。这意味着在测试集中，多层感知器分类器正确分类的图像数量占测试集中图像总数的86.81%。

总结来说，在这三种分类器中，随机森林分类器的准确率最高，为93.06%，多层感知器分类器的准确率最低，为86.81%。SVM分类器的准确率为90.97%，处于中间水平。

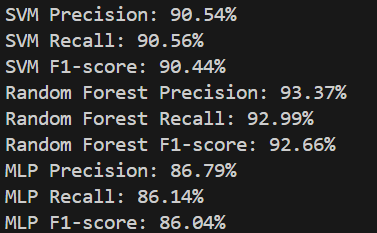
4.4.3 计算分类器指标



这段代码是使用Python的scikit-learn库来计算三个不同分类器（SVM、随机森林和多层感知器）在测试集上的性能指标。具体来说，它计算了每个分类器的精确度（Precision）、召回率（Recall）和F1分数（F1-score），并将结果打印出来。

首先，从scikit-learn库中导入precision\_score、recall\_score和f1\_score函数，这些函数用于计算性能指标。分别计算SVM、随机森林和多层感知器分类器的精确度、召回率和F1分数。这里使用了macro-averaging方法，它对每个类别的性能指标进行平均，从而得到整个分类器的性能指标。打印每个分类器的精确度、召回率和F1分数。

结果：



这段代码的运行结果给出了使用三种不同分类器（SVM、随机森林、多层感知器）在测试集上的性能指标。

- SVM Precision: 90.54%

SVM（支持向量机）分类器的精确度是90.54%。这意味着在测试集中，SVM分类器正确分类的图像数量占测试集中图像总数的90.54%。

- SVM Recall: 90.56%

SVM分类器的召回率是90.56%。这意味着在测试集中，SVM分类器正确分类的图像数量占实际类别为该类的图像总数的90.56%。

- SVM F1-score: 90.44%

SVM分类器的F1分数是90.44%。F1分数是精确度和召回率的调和平均值，用于衡量分类器的性能。

- Random Forest Precision: 93.37%

随机森林分类器的精确度是93.37%。这意味着在测试集中，随机森林分类器正确分类的图像数量占测试集中图像总数的93.37%。

- Random Forest Recall: 92.99%

随机森林分类器的召回率是92.99%。这意味着在测试集中，随机森林分类器正确分类的图像数量占实际类别为该类的图像总数的92.99%。

- Random Forest F1-score: 92.66%

随机森林分类器的F1分数是92.66%。F1分数是精确度和召回率的调和平均值，用于衡量分类器的性能。

- MLP Precision: 86.79%

多层感知器（MLP）分类器的精确度是86.79%。这意味着在测试集中，多层感知器分类器正确分类的图像数量占测试集中图像总数的86.79%。

- MLP Recall: 86.14%

多层感知器分类器的召回率是86.14%。这意味着在测试集中，多层感知器分类器正确分类的图像数量占实际类别为该类的图像总数的86.14%。

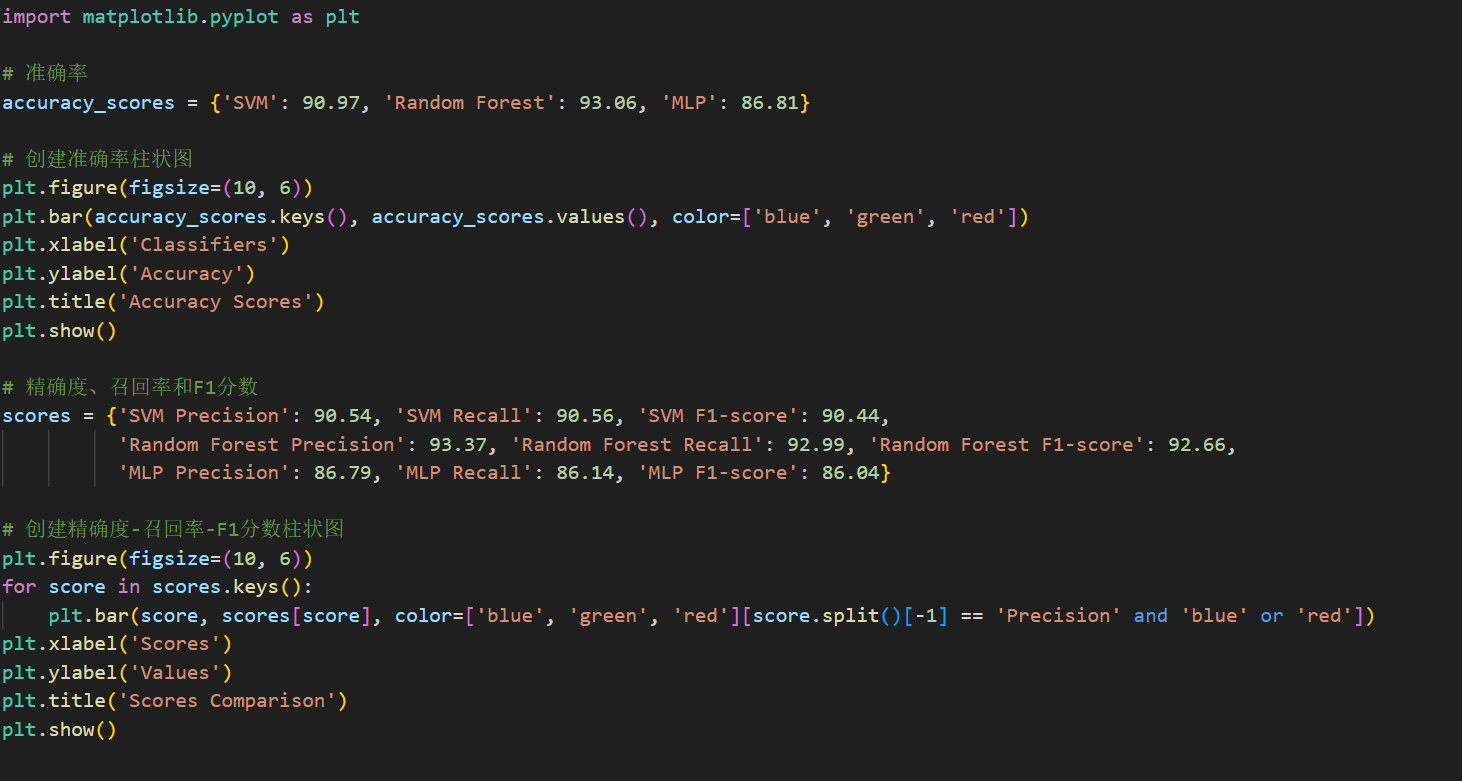
- MLP F1-score: 86.04%

多层感知器分类器的F1分数是86.04%。F1分数是精确度和召回率的调和平均值，用于衡量分类器的性能。

总结来说，在这三种分类器中，随机森林分类器的性能最好，其精确度、召回率和F1分数分别为93.37%、92.99%和92.66%。多层感知器分类器的性能最差，其精确度、召回率和F1分数分别为86.79%、86.14%和86.04%。SVM分类器的性能介于两者之间，其精确度、召回率和F1分数分别为90.54%、90.56%和90.44%。

# 5数据可视化

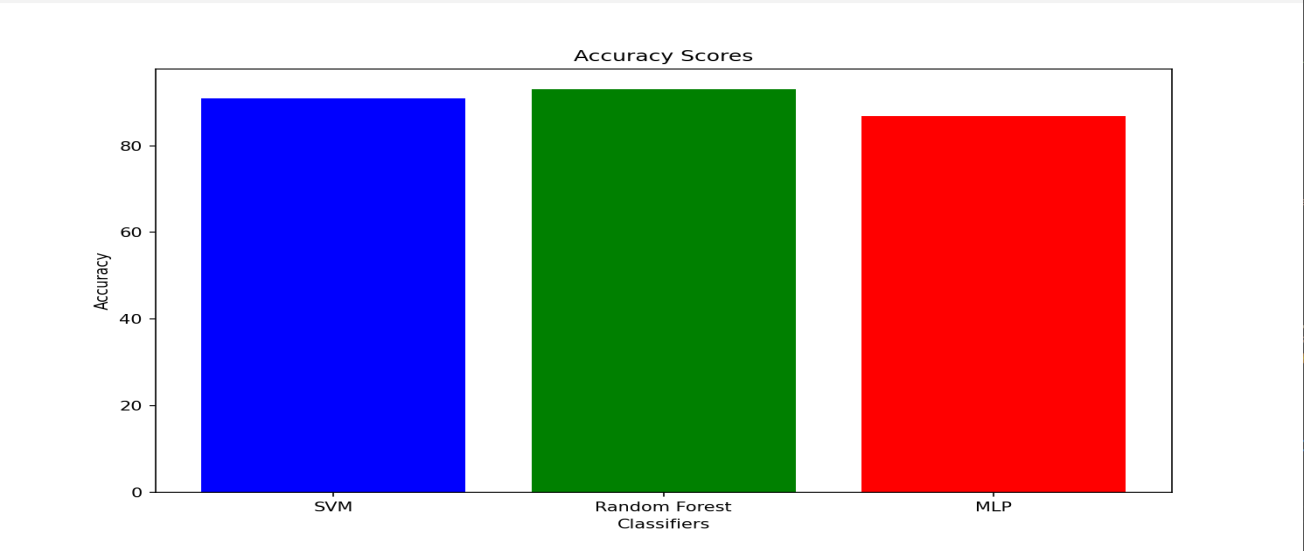
5.1代码部分



这段Python代码的目的是生成两个柱状图，分别显示不同分类器在准确率、精确度、召回率和F1分数方面的表现。第一个柱状图显示了三种分类器（SVM、随机森林和多层感知器）的准确率。x轴表示分类器，y轴表示准确率。从图表中可以看出，随机森林的准确率最高，为93.06%，而SVM的准确率最低，为90.97%。

第二个柱状图显示了三种分类器的精确度、召回率和F1分数。x轴表示指标，y轴表示值。从图表中可以看出，在精确度方面，随机森林的值最高，为93.37%，而SVM的值最低，为86.79%。在召回率方面，随机森林的值最高，为92.99%，而SVM的值最低，为86.14%。在F1分数方面，随机森林的值最高，为92.66%，而SVM的值最低，为86.04%。

可视化图表



# 6机器学习结论与建议

### 6.1 机器学习的结论

以上代码首先对图片数据进行了预处理，将图片转换为灰度图并resize到28x28大小。然后使用SVM、随机森林和多层感知器三种分类器对数据进行了训练和预测，并计算了它们的准确率、精确度、召回率和F1分数。最后，使用matplotlib库生成了准确率和精确度、召回率、F1分数的柱状图。

在准确率方面，随机森林分类器的表现最好，为93.06%，而多层感知器分类器的表现最差，为86.81%。支持向量机分类器介于两者之间，为90.97%。在精确度方面，随机森林分类器的表现最好，为93.37%，而多层感知器分类器的表现最差，为86.79%。支持向量机分类器介于两者之间，为90.54%。在召回率方面，随机森林分类器的表现最好，为92.99%，而多层感知器分类器的表现最差，为86.14%。支持向量机分类器介于两者之间，为90.56%。在F1分数方面，随机森林分类器的表现最好，为92.66%，而多层感知器分类器的表现最差，为86.04%。支持向量机分类器介于两者之间，为90.44%。

综上所述，在所评估的分类器中，随机森林分类器在各个方面都表现得最为出色，而多层感知器分类器在各个方面都表现得最为差劲。支持向量机分类器在准确率和F1分数方面表现较好，但在精确度和召回率方面表现一般。